



SEQUENCE LISTING

(1) GENERAL INFORMATION

(i) APPLICANT: Murphy, Patricia
White, Marga
Olson, Sheri
Yoshikawa, Matthew
Jackson, Geoffrey
Eskanderi, Tara
Schryer, Brenda
Park, Michael

(ii) TITLE OF THE INVENTION: NOVEL CODING SEQUENCE HAPLOTYPES
OF THE HUMAN BRCA2 GENE

(iii) NUMBER OF SEQUENCES: 133

(iv) CORRESPONDENCE ADDRESS:

- D |
(A) ADDRESSEE: Morgan, Lewis & Bockius LLP
(B) STREET: 1800 M St., NW
(C) CITY: Washington
(D) STATE: DC
(E) COUNTRY: USA
(F) ZIP: 20036

(v) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Diskette
(B) COMPUTER: IBM Compatible
(C) OPERATING SYSTEM: DOS
(D) SOFTWARE: FastSEQ for Windows Version 2.0

(vi) CURRENT APPLICATION DATA:

- (A) APPLICATION NUMBER: US 09/084,471
(B) FILING DATE: 22 May 1998
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

- (A) APPLICATION NUMBER: US 60/055,784
(B) FILING DATE: 15 August 1997

(viii) PRIOR APPLICATION DATA:

- (A) APPLICATION NUMBER: US 60/064,926
(B) FILING DATE: 07 November 1997

(ix) PRIOR APPLICATION DATA:

- (A) APPLICATION NUMBER: US 60/065,367
(B) FILING DATE: 12 November 1997

(x) PRIOR APPLICATION DATA:

- (A) APPLICATION NUMBER: US 09/071,715
(B) FILING DATE: 01 May 1998

(xi) ATTORNEY/AGENT INFORMATION:

- (A) NAME: Reid G. Adler
(B) REGISTRATION NUMBER: 30,988
(C) REFERENCE/DOCKET NUMBER: 44921-5058-01-US

(xii) TELECOMMUNICATION INFORMATION:

RECEIVED

OCT 25 2001

TECH CENTER 1600/2900

(A) TELEPHONE: 202-467-7000
(B) TELEFAX: 202-467-7258

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 105 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: exon
- (B) LOCATION: 26...75
- (D) OTHER INFORMATION: Exon 5

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

D /
cont

AGGGATTTCG TTTGTTTAT TTTAGTCCTG TTGTTCTACA ATGTACACAT GTAACACCAC 60
AAAGAGATAA GTCAGGTATG ATTAAAAACA ATGCTTTTA TTCTT 105

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 218 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: exon
- (B) LOCATION: 29...210
- (D) OTHER INFORMATION: Exon 15

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

TTTTTGCTAA GTATTTATTG TTTGATAGAT TTAATTACAA GTCTTCAGAA TGCCAGAGAT 60
ATACAGGATA TCGGAATTAA GAAGAAACAA AGGCAACGCG TCTTTCCACA GCCAGGCAGT 120
CTGTATCTTG CAAAAACATC CACTCTGCCT CGAATCTCTC TGAAAGCAGC AGTAGGAGGC 180
CAAGTTCCCT CTGCGTGTTC TCATAAACAG GTATGTGT 218

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 258 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: exon
- (B) LOCATION: 34...221
- (D) OTHER INFORMATION: Exon 16

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

tttttctttt ttgtgtgtgt ttatttgtg tagctgtata cgtatggcgt ttctaaacat 60

tgcataaaaa ttaacagcaa aaatgcagag tctttcagt ttcacactga agattatttt 120
 ggtaaagaaa gtttatggac tggaaaagga atacagttgg ctgatggtgg atggctata 180
 ccctccaatg atggaaaggc tggaaaagaa gaattttata ggtactctat gcaaaaagat 240
 tgtgtgttaa cttttatg 258

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10485 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 229...10482
- (D) OTHER INFORMATION: BRCA2 (OMI1)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

*D
D
cont*
 GGTGGCGCGA GCTTCTGAAA CTAGGCGGCA GAGGCGGAGC CGCTGTGGCA CTGCTGCGCC 60
 TCTGCTGCGC CTCGGGTGTC TTTTGGCGCG GTGGGTCGCC GCCGGGAGAA GCGTGAGGGG 120
 ACAGATTGTG GACCGGGCGCG GTTTTGTCA GCTTACTCCG GCCAAAAAAAG AACTGCACCT 180
 CTGGAGCGGA CTTATTTACC AAGCATTGGA GGAATATCGT AGGTAAAAA ATG CCT ATT 237
Met Pro Ile
1

 GGA TCC AAA GAG AGG CCA ACA TTT TTT GAA ATT TTT AAG ACA CGC TGC 285
 Gly Ser Lys Glu Arg Pro Thr Phe Phe Glu Ile Phe Lys Thr Arg Cys
5 10 15

 AAC AAA GCA GAT TTA GGA CCA ATA AGT CTT AAT TGG TTT GAA GAA CTT 333
 Asn Lys Ala Asp Leu Gly Pro Ile Ser Leu Asn Trp Phe Glu Glu Leu
20 25 30 35

 TCT TCA GAA GCT CCA CCC TAT AAT TCT GAA CCT GCA GAA GAA TCT GAA 381
 Ser Ser Glu Ala Pro Pro Tyr Asn Ser Glu Pro Ala Glu Glu Ser Glu
40 45 50

 CAT AAA AAC AAC AAT TAC GAA CCA AAC CTA TTT AAA ACT CCA CAA AGG 429
 His Lys Asn Asn Tyr Glu Pro Asn Leu Phe Lys Thr Pro Gln Arg
55 60 65

 AAA CCA TCT TAT AAT CAG CTG GCT TCA ACT CCA ATA ATA TTC AAA GAG 477
 Lys Pro Ser Tyr Asn Gln Leu Ala Ser Thr Pro Ile Ile Phe Lys Glu
70 75 80

 CAA GGG CTG ACT CTG CCG CTG TAC CAA TCT CCT GTA AAA GAA TTA GAT 525
 Gln Gly Leu Thr Leu Pro Leu Tyr Gln Ser Pro Val Lys Glu Leu Asp
85 90 95

 AAA TTC AAA TTA GAC TTA GGA AGG AAT GTT CCC AAT AGT AGA CAT AAA 573
 Lys Phe Lys Leu Asp Leu Gly Arg Asn Val Pro Asn Ser Arg His Lys
100 105 110 115

 AGT CTT CGC ACA GTG AAA ACT AAA ATG GAT CAA GCA GAT GAT GTT TCC 621
 Ser Leu Arg Thr Val Lys Thr Lys Met Asp Gln Ala Asp Asp Val Ser
120 125 130

D
C
M

| | |
|---|------|
| TGT CCA CTT CTA AAT TCT TGT CTT AGT GAA AGT CCT GTT GTT CTA CAA Cys Pro Leu Leu Asn Ser Cys Leu Ser Glu Ser Pro Val Val Leu Gln 135 140 145 | 669 |
| TGT ACA CAT GTA ACA CCA CAA AGA GAT AAG TCA GTG GTA TGT GGG AGT Cys Thr His Val Thr Pro Gln Arg Asp Lys Ser Val Val Cys Gly Ser 150 155 160 | 717 |
| TTG TTT CAT ACA CCA AAG TTT GTG AAG GGT CGT CAG ACA CCA AAA CAT Leu Phe His Thr Pro Lys Phe Val Lys Gly Arg Gln Thr Pro Lys His 165 170 175 | 765 |
| ATT TCT GAA AGT CTA GGA GCT GAG GTG GAT CCT GAT ATG TCT TGG TCA Ile Ser Glu Ser Leu Gly Ala Glu Val Asp Pro Asp Met Ser Trp Ser 180 185 190 195 | 813 |
| AGT TCT TTA GCT ACA CCA CCC ACC CTT AGT TCT ACT GTG CTC ATA GTC Ser Ser Leu Ala Thr Pro Pro Thr Leu Ser Ser Thr Val Leu Ile Val 200 205 210 | 861 |
| AGA AAT GAA GAA GCA TCT GAA ACT GTA TTT CCT CAT GAT ACT ACT GCT Arg Asn Glu Glu Ala Ser Glu Thr Val Phe Pro His Asp Thr Thr Ala 215 220 225 | 909 |
| AAT GTG AAA AGC TAT TTT TCC AAT CAT GAT GAA AGT CTG AAG AAA AAT Asn Val Lys Ser Tyr Phe Ser Asn His Asp Glu Ser Leu Lys Lys Asn 230 235 240 | 957 |
| GAT AGA TTT ATC GCT TCT GTG ACA GAC AGT GAA AAC ACA AAT CAA AGA Asp Arg Phe Ile Ala Ser Val Thr Asp Ser Glu Asn Thr Asn Gln Arg 245 250 255 | 1005 |
| GAA GCT GCA AGT CAT GGA TTT GGA AAA ACA TCA GGG AAT TCA TTT AAA Glu Ala Ala Ser His Gly Phe Gly Lys Thr Ser Gly Asn Ser Phe Lys 260 265 270 275 | 1053 |
| GTA AAT AGC TGC AAA GAC CAC ATT GGA AAG TCA ATG CCA AAT GTC CTA Val Asn Ser Cys Lys Asp His Ile Gly Lys Ser Met Pro Asn Val Leu 280 285 290 | 1101 |
| GAA GAT GAA GTA TAT GAA ACA GTT GTA GAT ACC TCT GAA GAA GAT AGT Glu Asp Glu Val Tyr Glu Thr Val Val Asp Thr Ser Glu Glu Asp Ser 295 300 305 | 1149 |
| TTT TCA TTA TGT TTT TCT AAA TGT AGA ACA AAA AAT CTA CAA AAA GTA Phe Ser Leu Cys Phe Ser Lys Cys Arg Thr Lys Asn Leu Gln Lys Val 310 315 320 | 1197 |
| AGA ACT AGC AAG ACT AGG AAA AAA ATT TTC CAT GAA GCA AAC GCT GAT Arg Thr Ser Lys Thr Arg Lys Lys Ile Phe His Glu Ala Asn Ala Asp 325 330 335 | 1245 |
| GAA TGT GAA AAA TCT AAA AAC CAA GTG AAA GAA AAA TAC TCA TTT GTA Glu Cys Glu Lys Ser Lys Asn Gln Val Lys Glu Lys Tyr Ser Phe Val 340 345 350 355 | 1293 |
| TCT GAA GTG GAA CCA AAT GAT ACT GAT CCA TTA GAT TCA AAT GTA GCA Ser Glu Val Glu Pro Asn Asp Thr Asp Pro Leu Asp Ser Asn Val Ala 360 365 370 | 1341 |
| CAT CAG AAG CCC TTT GAG AGT GGA AGT GAC AAA ATC TCC AAG GAA GTT | 1389 |

His Gln Lys Pro Phe Glu Ser Gly Ser Asp Lys Ile Ser Lys Glu Val
 375 380 385

GTA CCG TCT TTG GCC TGT GAA TGG TCT CAA CTA ACC CTT TCA GGT CTA 1437
 Val Pro Ser Leu Ala Cys Glu Trp Ser Gln Leu Thr Leu Ser Gly Leu
 390 395 400

AAT GGA GCC CAG ATG GAG AAA ATA CCC CTA TTG CAT ATT TCT TCA TGT 1485
 Asn Gly Ala Gln Met Glu Lys Ile Pro Leu Leu His Ile Ser Ser Cys
 405 410 415

GAC CAA AAT ATT TCA GAA AAA GAC CTA TTA GAC ACA GAG AAC AAA AGA 1533
 Asp Gln Asn Ile Ser Glu Lys Asp Leu Leu Asp Thr Glu Asn Lys Arg
 420 425 430 435

AAG AAA GAT TTT CTT ACT TCA GAG AAT TCT TTG CCA CGT ATT TCT AGC 1581
 Lys Lys Asp Phe Leu Thr Ser Glu Asn Ser Leu Pro Arg Ile Ser Ser
 440 445 450

CTA CCA AAA TCA GAG AAG CCA TTA AAT GAG GAA ACA GTG GTA AAT AAG 1629
 Leu Pro Lys Ser Glu Lys Pro Leu Asn Glu Glu Thr Val Val Asn Lys
 455 460 465

AGA GAT GAA GAG CAG CAT CTT GAA TCT CAT ACA GAC TGC ATT CTT GCA 1677
 Arg Asp Glu Glu Gln His Leu Glu Ser His Thr Asp Cys Ile Leu Ala
 470 475 480

GTA AAG CAG GCA ATA TCT GGA ACT TCT CCA GTG GCT TCT TCA TTT CAG 1725
 Val Lys Gln Ala Ile Ser Gly Thr Ser Pro Val Ala Ser Ser Phe Gln
 485 490 495

GGT ATC AAA AAG TCT ATA TTC AGA ATA AGA GAA TCA CCT AAA GAG ACT 1773
 Gly Ile Lys Lys Ser Ile Phe Arg Ile Arg Glu Ser Pro Lys Glu Thr
 500 505 510 515

TTC AAT GCA AGT TTT TCA GGT CAT ATG ACT GAT CCA AAC TTT AAA AAA 1821
 Phe Asn Ala Ser Phe Ser Gly His Met Thr Asp Pro Asn Phe Lys Lys
 520 525 530

GAA ACT GAA GCC TCT GAA AGT GGA CTG GAA ATA CAT ACT GTT TGC TCA 1869
 Glu Thr Glu Ala Ser Glu Ser Gly Leu Glu Ile His Thr Val Cys Ser
 535 540 545

CAG AAG GAG GAC TCC TTA TGT CCA AAT TTA ATT GAT AAT GGA AGC TGG 1917
 Gln Lys Glu Asp Ser Leu Cys Pro Asn Leu Ile Asp Asn Gly Ser Trp
 550 555 560

CCA GCC ACC ACC ACA CAG AAT TCT GTA GCT TTG AAG AAT GCA GGT TTA 1965
 Pro Ala Thr Thr Gln Asn Ser Val Ala Leu Lys Asn Ala Gly Leu
 565 570 575

ATA TCC ACT TTG AAA AAG AAA ACA AAT AAG TTT ATT TAT GCT ATA CAT 2013
 Ile Ser Thr Leu Lys Lys Thr Asn Lys Phe Ile Tyr Ala Ile His
 580 585 590 595

GAT GAA ACA TCT TAT AAA GGA AAA AAA ATA CCG AAA GAC CAA AAA TCA 2061
 Asp Glu Thr Ser Tyr Lys Gly Lys Lys Ile Pro Lys Asp Gln Lys Ser
 600 605 610

GAA CTA ATT AAC TGT TCA GCC CAG TTT GAA GCA AAT GCT TTT GAA GCA 2109
 Glu Leu Ile Asn Cys Ser Ala Gln Phe Glu Ala Asn Ala Phe Glu Ala

615 620 625

| | |
|--|------|
| CCA CTT ACA TTT GCA AAT GCT GAT TCA GGT TTA TTG CAT TCT TCT GTG Pro Leu Thr Phe Ala Asn Ala Asp Ser Gly Leu Leu His Ser Ser Val 630 635 640 | 2157 |
| AAA AGA AGC TGT TCA CAG AAT GAT TCT GAA GAA CCA ACT TTG TCC TTA Lys Arg Ser Cys Ser Gln Asn Asp Ser Glu Glu Pro Thr Leu Ser Leu 645 650 655 | 2205 |
| ACT AGC TCT TTT GGG ACA ATT CTG AGG AAA TGT TCT AGA AAT GAA ACA Thr Ser Ser Phe Gly Thr Ile Leu Arg Lys Cys Ser Arg Asn Glu Thr 660 665 670 675 | 2253 |
| TGT TCT AAT AAT ACA GTA ATC TCT CAG GAT CTT GAT TAT AAA GAA GCA Cys Ser Asn Asn Thr Val Ile Ser Gln Asp Leu Asp Tyr Lys Glu Ala 680 685 690 | 2301 |
| AAA TGT AAT AAG GAA AAA CTA CAG TTA TTT ATT ACC CCA GAA GCT GAT Lys Cys Asn Lys Glu Lys Leu Gln Leu Phe Ile Thr Pro Glu Ala Asp 695 700 705 | 2349 |
| TCT CTG TCA TGC CTG CAG GAA GGA CAG TGT GAA AAT GAT CCA AAA AGC Ser Leu Ser Cys Leu Gln Glu Gln Cys Glu Asn Asp Pro Lys Ser 710 715 720 | 2397 |
| AAA AAA GTT TCA GAT ATA AAA GAA GAG GTC TTG GCT GCA GCA TGT CAC Lys Lys Val Ser Asp Ile Lys Glu Glu Val Leu Ala Ala Ala Cys His 725 730 735 | 2445 |
| CCA GTA CAA CAT TCA AAA GTG GAA TAC AGT GAT ACT GAC TTT CAA TCC Pro Val Gln His Ser Lys Val Glu Tyr Ser Asp Thr Asp Phe Gln Ser 740 745 750 755 | 2493 |
| CAG AAA AGT CTT TTA TAT GAT CAT GAA AAT GCC AGC ACT CTT ATT TTA Gln Lys Ser Leu Leu Tyr Asp His Glu Asn Ala Ser Thr Leu Ile Leu 760 765 770 | 2541 |
| ACT CCT ACT TCC AAG GAT GTT CTG TCA AAC CTA GTC ATG ATT TCT AGA Thr Pro Thr Ser Lys Asp Val Leu Ser Asn Leu Val Met Ile Ser Arg 775 780 785 | 2589 |
| GGC AAA GAA TCA TAC AAA ATG TCA GAC AAG CTC AAA GGT AAC AAT TAT Gly Lys Glu Ser Tyr Lys Met Ser Asp Lys Leu Lys Gly Asn Asn Tyr 790 795 800 | 2637 |
| GAA TCT GAT GTT GAA TTA ACC AAA AAT ATT CCC ATG GAA AAG AAT CAA Glu Ser Asp Val Glu Leu Thr Lys Asn Ile Pro Met Glu Lys Asn Gln 805 810 815 | 2685 |
| GAT GTA TGT GCT TTA AAT GAA AAT TAT AAA AAC GTT GAG CTG TTG CCA Asp Val Cys Ala Leu Asn Glu Asn Tyr Lys Asn Val Glu Leu Leu Pro 820 825 830 835 | 2733 |
| CCT GAA AAA TAC ATG AGA GTA GCA TCA CCT TCA AGA AAG GTA CAA TTC Pro Glu Lys Tyr Met Arg Val Ala Ser Pro Ser Arg Lys Val Gln Phe 840 845 850 | 2781 |
| AAC CAA AAC ACA AAT CTA AGA GTA ATC CAA AAA AAT CAA GAA GAA ACT Asn Gln Asn Thr Asn Leu Arg Val Ile Gln Lys Asn Gln Glu Glu Thr 855 860 865 | 2829 |

ACT TCA ATT TCA AAA ATA ACT GTC AAT CCA GAC TCT GAA GAA CTT TTC 2877
 Thr Ser Ile Ser Lys Ile Thr Val Asn Pro Asp Ser Glu Glu Leu Phe
 870 875 880

 TCA GAC AAT GAG AAT AAT TTT GTC TTC CAA GTA GCT AAT GAA AGG AAT 2925
 Ser Asp Asn Glu Asn Asn Phe Val Phe Gln Val Ala Asn Glu Arg Asn
 885 890 895

 AAT CTT GCT TTA GGA AAT ACT AAG GAA CTT CAT GAA ACA GAC TTG ACT 2973
 Asn Leu Ala Leu Gly Asn Thr Lys Glu Leu His Glu Thr Asp Leu Thr
 900 905 910 915

 TGT GTA AAC GAA CCC ATT TTC AAG AAC TCT ACC ATG GTT TTA TAT GGA 3021
 Cys Val Asn Glu Pro Ile Phe Lys Asn Ser Thr Met Val Leu Tyr Gly
 920 925 930

 GAC ACA GGT GAT AAA CAA GCA ACC CAA GTG TCA ATT AAA AAA GAT TTG 3069
 Asp Thr Gly Asp Lys Gln Ala Thr Gln Val Ser Ile Lys Lys Asp Leu
 935 940 945

 GTT TAT GTT CTT GCA GAG GAG AAC AAA AAT AGT GTA AAG CAG CAT ATA 3117
 Val Tyr Val Leu Ala Glu Glu Asn Lys Asn Ser Val Lys Gln His Ile
 950 955 960

 AAA ATG ACT CTA GGT CAA GAT TTA AAA TCG GAC ATC TCC TTG AAT ATA 3165
 Lys Met Thr Leu Gly Gln Asp Leu Lys Ser Asp Ile Ser Leu Asn Ile
 965 970 975

 GAT AAA ATA CCA GAA AAA AAT AAT GAT TAC ATG AAC AAA TGG GCA GGA 3213
 Asp Lys Ile Pro Glu Lys Asn Asn Asp Tyr Met Asn Lys Trp Ala Gly
 980 985 990 995

 CTC TTA GGT CCA ATT TCA AAT CAC AGT TTT GGA GGT AGC TTC AGA ACA 3261
 Leu Leu Gly Pro Ile Ser Asn His Ser Phe Gly Gly Ser Phe Arg Thr
 1000 1005 1010

 GCT TCA AAT AAG GAA ATC AAG CTC TCT GAA CAT AAC ATT AAG AAG AGC 3309
 Ala Ser Asn Lys Glu Ile Lys Leu Ser Glu His Asn Ile Lys Ser
 1015 1020 1025

 AAA ATG TTC TTC AAA GAT ATT GAA GAA CAA TAT CCT ACT AGT TTA GCT 3357
 Lys Met Phe Phe Lys Asp Ile Glu Glu Gln Tyr Pro Thr Ser Leu Ala
 1030 1035 1040

 TGT GTT GAA ATT GTA AAT ACC TTG GCA TTA GAT AAT CAA AAG AAA CTG 3405
 Cys Val Glu Ile Val Asn Thr Leu Ala Leu Asp Asn Gln Lys Lys Leu
 1045 1050 1055

 AGC AAG CCT CAG TCA ATT AAT ACT GTA TCT GCA CAT TTA CAG AGT AGT 3453
 Ser Lys Pro Gln Ser Ile Asn Thr Val Ser Ala His Leu Gln Ser Ser
 1060 1065 1070 1075

 GTA GTT GTT TCT GAT TGT AAA AAT AGT CAT ATA ACC CCT CAG ATG TTA 3501
 Val Val Val Ser Asp Cys Lys Asn Ser His Ile Thr Pro Gln Met Leu
 1080 1085 1090

 TTT TCC AAG CAG GAT TTT AAT TCA AAC CAT AAT TTA ACA CCT AGC CAA 3549
 Phe Ser Lys Gln Asp Phe Asn Ser Asn His Asn Leu Thr Pro Ser Gln
 1095 1100 1105

D
cont

| | |
|---|------|
| AAG GCA GAA ATT ACA GAA CTT TCT ACT ATA TTA GAA GAA TCA GGA AGT Lys Ala Glu Ile Thr Glu Leu Ser Thr Ile Leu Glu Ser Gly Ser 1110 1115 1120 | 3597 |
| CAG TTT GAA TTT ACT CAG TTT AGA AAA CCA AGC TAC ATA TTG CAG AAG Gln Phe Glu Phe Thr Gln Phe Arg Lys Pro Ser Tyr Ile Leu Gln Lys 1125 1130 1135 | 3645 |
| AGT ACA TTT GAA GTG CCT GAA AAC CAG ATG ACT ATC TTA AAG ACC ACT Ser Thr Phe Glu Val Pro Glu Asn Gln Met Thr Ile Leu Lys Thr Thr 1140 1145 1150 1155 | 3693 |
| TCT GAG GAA TGC AGA GAT GCT GAT CTT CAT GTC ATA ATG AAT GCC CCA Ser Glu Glu Cys Arg Asp Ala Asp Leu His Val Ile Met Asn Ala Pro 1160 1165 1170 | 3741 |
| TCG ATT GGT CAG GTA GAC AGC AGC AAG CAA TTT GAA GGT ACA GTT GAA Ser Ile Gly Gln Val Asp Ser Ser Lys Gln Phe Glu Gly Thr Val Glu 1175 1180 1185 | 3789 |
| ATT AAA CGG AAG TTT GCT GGC CTG TTG AAA AAT GAC TGT AAC AAA AGT Ile Lys Arg Lys Phe Ala Gly Leu Leu Lys Asn Asp Cys Asn Lys Ser 1190 1195 1200 | 3837 |
| GCT TCT GGT TAT TTA ACA GAT GAA AAT GAA GTG GGG TTT AGG GGC TTT Ala Ser Gly Tyr Leu Thr Asp Glu Asn Glu Val Gly Phe Arg Gly Phe 1205 1210 1215 | 3885 |
| TAT TCT GCT CAT GGC ACA AAA CTG AAT GTT TCT ACT GAA GCT CTG CAA Tyr Ser Ala His Gly Thr Lys Leu Asn Val Ser Thr Glu Ala Leu Gln 1220 1225 1230 1235 | 3933 |
| AAA GCT GTG AAA CTG TTT AGT GAT ATT GAG AAT ATT AGT GAG GAA ACT Lys Ala Val Lys Leu Phe Ser Asp Ile Glu Asn Ile Ser Glu Glu Thr 1240 1245 1250 | 3981 |
| TCT GCA GAG GTA CAT CCA ATA AGT TTA TCT TCA AGT AAA TGT CAT GAT Ser Ala Glu Val His Pro Ile Ser Leu Ser Ser Lys Cys His Asp 1255 1260 1265 | 4029 |
| TCT GTT GTT TCA ATG TTT AAG ATA GAA AAT CAT AAT GAT AAA ACT GTA Ser Val Val Ser Met Phe Lys Ile Glu Asn His Asn Asp Lys Thr Val 1270 1275 1280 | 4077 |
| AGT GAA AAA AAT AAT AAA TGC CAA CTG ATA TTA CAA AAT AAT ATT GAA Ser Glu Lys Asn Asn Lys Cys Gln Leu Ile Leu Gln Asn Asn Ile Glu 1285 1290 1295 | 4125 |
| ATG ACT ACT GGC ACT TTT GTT GAA GAA ATT ACT GAA AAT TAC AAG AGA Met Thr Thr Gly Thr Phe Val Glu Glu Ile Thr Glu Asn Tyr Lys Arg 1300 1305 1310 1315 | 4173 |
| AAT ACT GAA AAT GAA GAT AAC AAA TAT ACT GCT GCC AGT AGA AAT TCT Asn Thr Glu Asn Glu Asp Asn Lys Tyr Thr Ala Ala Ser Arg Asn Ser 1320 1325 1330 | 4221 |
| CAT AAC TTA GAA TTT GAT GGC AGT GAT TCA AGT AAA AAT GAT ACT GTT His Asn Leu Glu Phe Asp Gly Ser Asp Ser Ser Lys Asn Asp Thr Val 1335 1340 1345 | 4269 |
| TGT ATT CAT AAA GAT GAA ACG GAC TTG CTA TTT ACT GAT CAG CAC AAC | 4317 |

D
1
cont

| | | | |
|---|------|------|------|
| Cys Ile His Lys Asp Glu Thr Asp Leu Leu Phe Thr Asp Gln His Asn | | | |
| 1350 | 1355 | 1360 | |
| ATA TGT CTT AAA TTA TCT GGC CAG TTT ATG AAG GAG GGA AAC ACT CAG | | | 4365 |
| Ile Cys Leu Lys Leu Ser Gly Gln Phe Met Lys Glu Gly Asn Thr Gln | | | |
| 1365 | 1370 | 1375 | |
| ATT AAA GAA GAT TTG TCA GAT TTA ACT TTT TTG GAA GTT GCG AAA GCT | | | 4413 |
| Ile Lys Glu Asp Leu Ser Asp Leu Thr Phe Leu Glu Val Ala Lys Ala | | | |
| 1380 | 1385 | 1390 | 1395 |
| CAA GAA GCA TGT CAT GGT AAT ACT TCA AAT AAA GAA CAG TTA ACT GCT | | | 4461 |
| Gln Glu Ala Cys His Gly Asn Thr Ser Asn Lys Glu Gln Leu Thr Ala | | | |
| 1400 | 1405 | 1410 | |
| ACT AAA ACG GAG CAA AAT ATA AAA GAT TTT GAG ACT TCT GAT ACA TTT | | | 4509 |
| Thr Lys Thr Glu Gln Asn Ile Lys Asp Phe Glu Thr Ser Asp Thr Phe | | | |
| 1415 | 1420 | 1425 | |
| TTT CAG ACT GCA AGT GGG AAA AAT ATT AGT GTC GCC AAA GAG TCA TTT | | | 4557 |
| Phe Gln Thr Ala Ser Gly Lys Asn Ile Ser Val Ala Lys Glu Ser Phe | | | |
| 1430 | 1435 | 1440 | |
| AAT AAA ATT GTA AAT TTC TTT GAT CAG AAA CCA GAA GAA TTG CAT AAC | | | 4605 |
| Asn Lys Ile Val Asn Phe Phe Asp Gln Lys Pro Glu Glu Leu His Asn | | | |
| 1445 | 1450 | 1455 | |
| TTT TCC TTA AAT TCT GAA TTA CAT TCT GAC ATA AGA AAG AAC AAA ATG | | | 4653 |
| Phe Ser Leu Asn Ser Glu Leu His Ser Asp Ile Arg Lys Asn Lys Met | | | |
| 1460 | 1465 | 1470 | 1475 |
| GAC ATT CTA AGT TAT GAG GAA ACA GAC ATA GTT AAA CAC AAA ATA CTG | | | 4701 |
| Asp Ile Leu Ser Tyr Glu Glu Thr Asp Ile Val Lys His Lys Ile Leu | | | |
| 1480 | 1485 | 1490 | |
| AAA GAA AGT GTC CCA GTT GGT ACT GGA AAT CAA CTA GTG ACC TTC CAG | | | 4749 |
| Lys Glu Ser Val Pro Val Gly Thr Gly Asn Gln Leu Val Thr Phe Gln | | | |
| 1495 | 1500 | 1505 | |
| GGA CAA CCC GAA CGT GAT GAA AAG ATC AAA GAA CCT ACT CTG TTG GGT | | | 4797 |
| Gly Gln Pro Glu Arg Asp Glu Lys Ile Lys Glu Pro Thr Leu Leu Gly | | | |
| 1510 | 1515 | 1520 | |
| TTT CAT ACA GCT AGC GGG AAA AAA GTT AAA ATT GCA AAG GAA TCT TTG | | | 4845 |
| Phe His Thr Ala Ser Gly Lys Lys Val Lys Ile Ala Lys Glu Ser Leu | | | |
| 1525 | 1530 | 1535 | |
| GAC AAA GTG AAA AAC CTT TTT GAT GAA AAA GAG CAA GGT ACT AGT GAA | | | 4893 |
| Asp Lys Val Lys Asn Leu Phe Asp Glu Lys Glu Gln Gly Thr Ser Glu | | | |
| 1540 | 1545 | 1550 | 1555 |
| ATC ACC AGT TTT AGC CAT CAA TGG GCA AAG ACC CTA AAG TAC AGA GAG | | | 4941 |
| Ile Thr Ser Phe Ser His Gln Trp Ala Lys Thr Leu Lys Tyr Arg Glu | | | |
| 1560 | 1565 | 1570 | |
| GCC TGT AAA GAC CTT GAA TTA GCA TGT GAG ACC ATT GAG ATC ACA GCT | | | 4989 |
| Ala Cys Lys Asp Leu Glu Leu Ala Cys Glu Thr Ile Glu Ile Thr Ala | | | |
| 1575 | 1580 | 1585 | |
| GCC CCA AAG TGT AAA GAA ATG CAG AAT TCT CTC AAT AAT GAT AAA AAC | | | 5037 |
| Ala Pro Lys Cys Lys Glu Met Gln Asn Ser Leu Asn Asn Asp Lys Asn | | | |

1590 1595 1600

| | | | |
|--|------|------|------|
| CTT GTT TCT ATT GAG ACT GTG GTG CCA CCT AAG CTC TTA AGT GAT AAT Leu Val Ser Ile Glu Thr Val Val Pro Pro Lys Leu Leu Ser Asp Asn 1605 | 1610 | 1615 | 5085 |
| TTA TGT AGA CAA ACT GAA AAT CTC AAA ACA TCA AAA AGT ATC TTT TTG Leu Cys Arg Gln Thr Glu Asn Leu Lys Thr Ser Lys Ser Ile Phe Leu 1620 | 1625 | 1630 | 1635 |
| AAA GTT AAA GTA CAT GAA AAT GTA GAA AAA GAA ACA GCA AAA AGT CCT Lys Val Lys Val His Glu Asn Val Glu Lys Glu Thr Ala Lys Ser Pro 1640 | 1645 | 1650 | 5181 |
| GCA ACT TGT TAC ACA AAT CAG TCC CCT TAT TCA GTC ATT GAA AAT TCA Ala Thr Cys Tyr Thr Asn Gln Ser Pro Tyr Ser Val Ile Glu Asn Ser 1655 | 1660 | 1665 | 5229 |
| GCC TTA GCT TTT TAC ACA AGT TGT AGT AGA AAA ACT TCT GTG AGT CAG Ala Leu Ala Phe Tyr Thr Ser Cys Ser Arg Lys Thr Ser Val Ser Gln 1670 | 1675 | 1680 | 5277 |
| ACT TCA TTA CTT GAA GCA AAA AAA TGG CTT AGA GAA GGA ATA TTT GAT Thr Ser Leu Leu Glu Ala Lys Lys Trp Leu Arg Glu Gly Ile Phe Asp 1685 | 1690 | 1695 | 5325 |
| GGT CAA CCA GAA AGA ATA AAT ACT GCA GAT TAT GTA GGA AAT TAT TTG Gly Gln Pro Glu Arg Ile Asn Thr Ala Asp Tyr Val Gly Asn Tyr Leu 1700 | 1705 | 1710 | 1715 |
| TAT GAA AAT AAT TCA AAC AGT ACT ATA GCT GAA AAT GAC AAA AAT CAT Tyr Glu Asn Asn Ser Asn Ser Thr Ile Ala Glu Asn Asp Lys Asn His 1720 | 1725 | 1730 | 5421 |
| CTC TCC GAA AAA CAA GAT ACT TAT TTA AGT AAC AGT AGC ATG TCT AAC Leu Ser Glu Lys Gln Asp Thr Tyr Leu Ser Asn Ser Met Ser Asn 1735 | 1740 | 1745 | 5469 |
| AGC TAT TCC TAC CAT TCT GAT GAG GTA TAT AAT GAT TCA GGA TAT CTC Ser Tyr Ser Tyr His Ser Asp Glu Val Tyr Asn Asp Ser Gly Tyr Leu 1750 | 1755 | 1760 | 5517 |
| TCA AAA AAT AAA CTT GAT TCT GGT ATT GAG CCA GTA TTG AAG AAT GTT Ser Lys Asn Lys Leu Asp Ser Gly Ile Glu Pro Val Leu Lys Asn Val 1765 | 1770 | 1775 | 5565 |
| GAA GAT CAA AAA AAC ACT AGT TTT TCC AAA GTA ATA TCC AAT GTA AAA Glu Asp Gln Lys Asn Thr Ser Phe Ser Lys Val Ile Ser Asn Val Lys 1780 | 1785 | 1790 | 1795 |
| GAT GCA AAT GCA TAC CCA CAA ACT GTA AAT GAA GAT ATT TGC GTT GAG Asp Ala Asn Ala Tyr Pro Gln Thr Val Asn Glu Asp Ile Cys Val Glu 1800 | 1805 | 1810 | 5661 |
| GAA CTT GTG ACT AGC TCT TCA CCC TGC AAA AAT AAA AAT GCA GCC ATT Glu Leu Val Thr Ser Ser Pro Cys Lys Asn Lys Asn Ala Ala Ile 1815 | 1820 | 1825 | 5709 |
| AAA TTG TCC ATA TCT AAT AGT AAT AAT TTT GAG GTA GGG CCA CCT GCA Lys Leu Ser Ile Ser Asn Ser Asn Asn Phe Glu Val Gly Pro Pro Ala 1830 | 1835 | 1840 | 5757 |

| | |
|---|------|
| TTT AGG ATA GCC AGT GGT AAA ATC GTT TGT GTT TCA CAT GAA ACA ATT Phe Arg Ile Ala Ser Gly Lys Ile Val Cys Val Ser His Glu Thr Ile 1845 1850 1855 | 5805 |
| AAA AAA GTG AAA GAC ATA TTT ACA GAC AGT TTC AGT AAA GTA ATT AAG Lys Lys Val Lys Asp Ile Phe Thr Asp Ser Phe Ser Lys Val Ile Lys 1860 1865 1870 1875 | 5853 |
| GAA AAC AAC GAG AAT AAA TCA AAA ATT TGC CAA ACG AAA ATT ATG GCA Glu Asn Asn Glu Asn Lys Ser Lys Ile Cys Gln Thr Lys Ile Met Ala 1880 1885 1890 | 5901 |
| GGT TGT TAC GAG GCA TTG GAT GAT TCA GAG GAT ATT CTT CAT AAC TCT Gly Cys Tyr Glu Ala Leu Asp Asp Ser Glu Asp Ile Leu His Asn Ser 1895 1900 1905 | 5949 |
| D CON | |
| CTA GAT AAT GAT GAA TGT AGC ACG CAT TCA CAT AAG GTT TTT GCT GAC Leu Asp Asn Asp Glu Cys Ser Thr His Ser His Lys Val Phe Ala Asp 1910 1915 1920 | 5997 |
| ATT CAG AGT GAA GAA ATT TTA CAA CAT AAC CAA AAT ATG TCT GGA TTG Ile Gln Ser Glu Glu Ile Leu Gln His Asn Gln Asn Met Ser Gly Leu 1925 1930 1935 | 6045 |
| GAG AAA GTT TCT AAA ATA TCA CCT TGT GAT GTT AGT TTG GAA ACT TCA Glu Lys Val Ser Lys Ile Ser Pro Cys Asp Val Ser Leu Glu Thr Ser 1940 1945 1950 1955 | 6093 |
| GAT ATA TGT AAA TGT AGT ATA GGG AAG CTT CAT AAG TCA GTC TCA TCT Asp Ile Cys Lys Cys Ser Ile Gly Lys Leu His Lys Ser Val Ser Ser 1960 1965 1970 | 6141 |
| GCA AAT ACT TGT GGG ATT TTT AGC ACA GCA AGT GGA AAA TCT GTC CAG Ala Asn Thr Cys Gly Ile Phe Ser Thr Ala Ser Gly Lys Ser Val Gln 1975 1980 1985 | 6189 |
| GTA TCA GAT GCT TCA TTA CAA AAC GCA AGA CAA GTG TTT TCT GAA ATA Val Ser Asp Ala Ser Leu Gln Asn Ala Arg Gln Val Phe Ser Glu Ile 1990 1995 2000 | 6237 |
| GAA GAT AGT ACC AAG CAA GTC TTT TCC AAA GTA TTG TTT AAA AGT AAC Glu Asp Ser Thr Lys Gln Val Phe Ser Lys Val Leu Phe Lys Ser Asn 2005 2010 2015 | 6285 |
| GAA CAT TCA GAC CAG CTC ACA AGA GAA GAA AAT ACT GCT ATA CGT ACT Glu His Ser Asp Gln Leu Thr Arg Glu Glu Asn Thr Ala Ile Arg Thr 2020 2025 2030 2035 | 6333 |
| CCA GAA CAT TTA ATA TCC CAA AAA GGC TTT TCA TAT AAT GTG GTA AAT Pro Glu His Leu Ile Ser Gln Lys Gly Phe Ser Tyr Asn Val Val Asn 2040 2045 2050 | 6381 |
| TCA TCT GCT TTC TCT GGA TTT AGT ACA GCA AGT GGA AAG CAA GTT TCC Ser Ser Ala Phe Ser Gly Phe Ser Thr Ala Ser Gly Lys Gln Val Ser 2055 2060 2065 | 6429 |
| ATT TTA GAA AGT TCC TTA CAC AAA GTT AAG GGA GTG TTA GAG GAA TTT Ile Leu Glu Ser Ser Leu His Lys Val Lys Gly Val Leu Glu Glu Phe 2070 2075 2080 | 6477 |

D
X
on

| | |
|---|------|
| GAT TTA ATC AGA ACT GAG CAT AGT CTT CAC TAT TCA CCT ACG TCT AGA Asp Leu Ile Arg Thr Glu His Ser Leu His Tyr Ser Pro Thr Ser Arg 2085 2090 2095 | 6525 |
| CAA AAT GTA TCA AAA ATA CTT CCT CGT GTT GAT AAG AGA AAC CCA GAG Gln Asn Val Ser Lys Ile Leu Pro Arg Val Asp Lys Arg Asn Pro Glu 2100 2105 2110 2115 | 6573 |
| CAC TGT GTA AAC TCA GAA ATG GAA AAA ACC TGC AGT AAA GAA TTT AAA His Cys Val Asn Ser Glu Met Glu Lys Thr Cys Ser Lys Glu Phe Lys 2120 2125 2130 | 6621 |
| TTA TCA AAT AAC TTA AAT GTT GAA GGT GGT TCT TCA GAA AAT AAT CAC Leu Ser Asn Asn Leu Asn Val Glu Gly Gly Ser Ser Glu Asn Asn His 2135 2140 2145 | 6669 |
| TCT ATT AAA GTT TCT CCA TAT CTC TCT CAA TTT CAA CAA GAC AAA CAA Ser Ile Lys Val Ser Pro Tyr Leu Ser Gln Phe Gln Gln Asp Lys Gln 2150 2155 2160 | 6717 |
| CAG TTG GTA TTA GGA ACC AAA GTC TCA CTT GTT GAG AAC ATT CAT GTT Gln Leu Val Leu Gly Thr Lys Val Ser Leu Val Glu Asn Ile His Val 2165 2170 2175 | 6765 |
| TTG GGA AAA GAA CAG GCT TCA CCT AAA AAC GTA AAA ATG GAA ATT GGT Leu Gly Lys Glu Gln Ala Ser Pro Lys Asn Val Lys Met Glu Ile Gly 2180 2185 2190 2195 | 6813 |
| AAA ACT GAA ACT TTT TCT GAT GTT CCT GTG AAA ACA AAT ATA GAA GTT Lys Thr Glu Thr Phe Ser Asp Val Pro Val Lys Thr Asn Ile Glu Val 2200 2205 2210 | 6861 |
| TGT TCT ACT TAC TCC AAA GAT TCA GAA AAC TAC TTT GAA ACA GAA GCA Cys Ser Thr Tyr Ser Lys Asp Ser Glu Asn Tyr Phe Glu Thr Glu Ala 2215 2220 2225 | 6909 |
| GTA GAA ATT GCT AAA GCT TTT ATG GAA GAT GAT GAA CTG ACA GAT TCT Val Glu Ile Ala Lys Ala Phe Met Glu Asp Asp Glu Leu Thr Asp Ser 2230 2235 2240 | 6957 |
| AAA CTG CCA AGT CAT GCC ACA CAT TCT CTT TTT ACA TGT CCC GAA AAT Lys Leu Pro Ser His Ala Thr His Ser Leu Phe Thr Cys Pro Glu Asn 2245 2250 2255 | 7005 |
| GAG GAA ATG GTT TTG TCA AAT TCA AGA ATT GGA AAA AGA AGA GGA GAG Glu Glu Met Val Leu Ser Asn Ser Arg Ile Gly Lys Arg Arg Gly Glu 2260 2265 2270 2275 | 7053 |
| CCC CTT ATC TTA GTG GGA GAA CCC TCA ATC AAA AGA AAC TTA TTA AAT Pro Leu Ile Leu Val Gly Glu Pro Ser Ile Lys Arg Asn Leu Leu Asn 2280 2285 2290 | 7101 |
| GAA TTT GAC AGG ATA ATA GAA AAT CAA GAA AAA TCC TTA AAG GCT TCA Glu Phe Asp Arg Ile Ile Glu Asn Gln Glu Lys Ser Leu Lys Ala Ser 2295 2300 2305 | 7149 |
| AAA AGC ACT CCA GAT GGC ACA ATA AAA GAT CGA AGA TTG TTT ATG CAT Lys Ser Thr Pro Asp Gly Thr Ile Lys Asp Arg Arg Leu Phe Met His 2310 2315 2320 | 7197 |
| CAT GTT TCT TTA GAG CCG ATT ACC TGT GTA CCC TTT CGC ACA ACT AAG | 7245 |

D
cont

| | | | |
|---|------|------|------|
| His Val Ser Leu Glu Pro Ile Thr Cys Val Pro Phe Arg Thr Thr Lys | | | |
| 2325 | 2330 | 2335 | |
| GAA CGT CAA GAG ATA CAG AAT CCA AAT TTT ACC GCA CCT GGT CAA GAA | | | 7293 |
| Glu Arg Gln Glu Ile Gln Asn Pro Asn Phe Thr Ala Pro Gly Gln Glu | | | |
| 2340 | 2345 | 2350 | 2355 |
| TTT CTG TCT AAA TCT CAT TTG TAT GAA CAT CTG ACT TTG GAA AAA TCT | | | 7341 |
| Phe Leu Ser Lys Ser His Leu Tyr Glu His Leu Thr Leu Glu Lys Ser | | | |
| 2360 | 2365 | 2370 | |
| TCA AGC AAT TTA GCA GTT TCA GGA CAT CCA TTT TAT CAA GTT TCT GCT | | | 7389 |
| Ser Ser Asn Leu Ala Val Ser Gly His Pro Phe Tyr Gln Val Ser Ala | | | |
| 2375 | 2380 | 2385 | |
| ACA AGA AAT GAA AAA ATG AGA CAC TTG ATT ACT ACA GGC AGA CCA ACC | | | 7437 |
| Thr Arg Asn Glu Lys Met Arg His Leu Ile Thr Thr Gly Arg Pro Thr | | | |
| 2390 | 2395 | 2400 | |
| AAA GTC TTT GTT CCA CCT TTT AAA ACT AAA TCA CAT TTT CAC AGA GTT | | | 7485 |
| Lys Val Phe Val Pro Pro Phe Lys Thr Lys Ser His Phe His Arg Val | | | |
| 2405 | 2410 | 2415 | |
| GAA CAG TGT GTT AGG AAT ATT AAC TTG GAG GAA AAC AGA CAA AAG CAA | | | 7533 |
| Glu Gln Cys Val Arg Asn Ile Asn Leu Glu Glu Asn Arg Gln Lys Gln | | | |
| 2420 | 2425 | 2430 | 2435 |
| AAC ATT GAT GGA CAT GGC TCT GAT GAT AGT AAA AAT AAG ATT AAT GAC | | | 7581 |
| Asn Ile Asp Gly His Gly Ser Asp Asp Ser Lys Asn Lys Ile Asn Asp | | | |
| 2440 | 2445 | 2450 | |
| AAT GAG ATT CAT CAG TTT AAC AAA AAC AAC TCC AAT CAA GCA GCA GCT | | | 7629 |
| Asn Glu Ile His Gln Phe Asn Lys Asn Asn Ser Asn Gln Ala Ala Ala | | | |
| 2455 | 2460 | 2465 | |
| GTA ACT TTC ACA AAG TGT GAA GAA CCT TTA GAT TTA ATT ACA AGT | | | 7677 |
| Val Thr Phe Thr Lys Cys Glu Glu Pro Leu Asp Leu Ile Thr Ser | | | |
| 2470 | 2475 | 2480 | |
| CTT CAG AAT GCC AGA GAT ATA CAG GAT ATG CGA ATT AAG AAG AAA CAA | | | 7725 |
| Leu Gln Asn Ala Arg Asp Ile Gln Asp Met Arg Ile Lys Lys Gln | | | |
| 2485 | 2490 | 2495 | |
| AGG CAA CGC GTC TTT CCA CAG CCA GGC AGT CTG TAT CTT GCA AAA ACA | | | 7773 |
| Arg Gln Arg Val Phe Pro Gln Pro Gly Ser Leu Tyr Leu Ala Lys Thr | | | |
| 2500 | 2505 | 2510 | 2515 |
| TCC ACT CTG CCT CGA ATC TCT CTG AAA GCA GCA GTA GGA GGC CAA GTT | | | 7821 |
| Ser Thr Leu Pro Arg Ile Ser Leu Lys Ala Ala Val Gly Gly Gln Val | | | |
| 2520 | 2525 | 2530 | |
| CCC TCT GCG TGT TCT CAT AAA CAG CTG TAT ACG TAT GGC GTT TCT AAA | | | 7869 |
| Pro Ser Ala Cys Ser His Lys Gln Leu Tyr Thr Tyr Gly Val Ser Lys | | | |
| 2535 | 2540 | 2545 | |
| CAT TGC ATA AAA ATT AAC AGC AAA AAT GCA GAG TCT TTT CAG TTT CAC | | | 7917 |
| His Cys Ile Lys Ile Asn Ser Lys Asn Ala Glu Ser Phe Gln Phe His | | | |
| 2550 | 2555 | 2560 | |
| ACT GAA GAT TAT TTT GGT AAG GAA AGT TTA TGG ACT GGA AAA GGA ATA | | | 7965 |
| Thr Glu Asp Tyr Phe Gly Lys Glu Ser Leu Trp Thr Gly Lys Gly Ile | | | |

D
cont

| | | | |
|---|------|------|------|
| 2565 | 2570 | 2575 | |
| CAG TTG GCT GAT GGT GGA TGG CTC ATA CCC TCC AAT GAT GGA AAG GCT Gln Leu Ala Asp Gly Gly Trp Leu Ile Pro Ser Asn Asp Gly Lys Ala 2580 2585 2590 2595 | | | 8013 |
| GGA AAA GAA GAA TTT TAT AGG GCT CTG TGT GAC ACT CCA GGT GTG GAT Gly Lys Glu Glu Phe Tyr Arg Ala Leu Cys Asp Thr Pro Gly Val Asp 2600 2605 2610 | | | 8061 |
| CCA AAG CTT ATT TCT AGA ATT TGG GTT TAT AAT CAC TAT AGA TGG ATC Pro Lys Leu Ile Ser Arg Ile Trp Val Tyr Asn His Tyr Arg Trp Ile 2615 2620 2625 | | | 8109 |
| ATA TGG AAA CTG GCA GCT ATG GAA TGT GCC TTT CCT AAG GAA TTT GCT Ile Trp Lys Leu Ala Ala Met Glu Cys Ala Phe Pro Lys Glu Phe Ala 2630 2635 2640 | | | 8157 |
| AAT AGA TGC CTA AGC CCA GAA AGG GTG CTT CTT CAA CTA AAA TAC AGA Asn Arg Cys Leu Ser Pro Glu Arg Val Leu Leu Gln Leu Lys Tyr Arg 2645 2650 2655 | | | 8205 |
| TAT GAT ACG GAA ATT GAT AGA AGC AGA AGA TCG GCT ATA AAA AAG ATA Tyr Asp Thr Glu Ile Asp Arg Ser Arg Ser Ala Ile Lys Lys Ile 2660 2665 2670 2675 | | | 8253 |
| ATG GAA AGG GAT GAC ACA GCT GCA AAA ACA CTT GTT CTC TGT GTT TCT Met Glu Arg Asp Asp Thr Ala Ala Lys Thr Leu Val Leu Cys Val Ser 2680 2685 2690 | | | 8301 |
| GAC ATA ATT TCA TTG AGC GCA AAT ATA TCT GAA ACT TCT AGC AAT AAA Asp Ile Ile Ser Leu Ser Ala Asn Ile Ser Glu Thr Ser Ser Asn Lys 2695 2700 2705 | | | 8349 |
| ACT AGT AGT GCA GAT ACC CAA AAA GTG GCC ATT ATT GAA CTT ACA GAT Thr Ser Ser Ala Asp Thr Gln Lys Val Ala Ile Ile Glu Leu Thr Asp 2710 2715 2720 | | | 8397 |
| GGG TGG TAT GCT GTT AAG GCC CAG TTA GAT CCT CCC CTC TTA GCT GTC Gly Trp Tyr Ala Val Lys Ala Gln Leu Asp Pro Pro Leu Leu Ala Val 2725 2730 2735 | | | 8445 |
| TTA AAG AAT GGC AGA CTG ACA GTT GGT CAG AAG ATT ATT CTT CAT GGA Leu Lys Asn Gly Arg Leu Thr Val Gly Gln Lys Ile Ile Leu His Gly 2740 2745 2750 2755 | | | 8493 |
| GCA GAA CTG GTG GGC TCT CCT GAT GCC TGT ACA CCT CTT GAA GCC CCA Ala Glu Leu Val Gly Ser Pro Asp Ala Cys Thr Pro Leu Glu Ala Pro 2760 2765 2770 | | | 8541 |
| GAA TCT CTT ATG TTA AAG ATT TCT GCT AAC AGT ACT CGG CCT GCT CGC Glu Ser Leu Met Leu Lys Ile Ser Ala Asn Ser Thr Arg Pro Ala Arg 2775 2780 2785 | | | 8589 |
| TGG TAT ACC AAA CTT GGA TTC TTT CCT GAC CCT AGA CCT TTT CCT CTG Trp Tyr Thr Lys Leu Gly Phe Phe Pro Asp Pro Arg Pro Phe Pro Leu 2790 2795 2800 | | | 8637 |
| CCC TTA TCA TCG CTT TTC AGT GAT GGA GGA AAT GTT GGT TGT GTT GAT Pro Leu Ser Ser Leu Phe Ser Asp Gly Gly Asn Val Gly Cys Val Asp 2805 2810 2815 | | | 8685 |

GTA ATT ATT CAA AGA GCA TAC CCT ATA CAG TGG ATG GAG AAG ACA TCA 8733
 Val Ile Ile Gln Arg Ala Tyr Pro Ile Gln Trp Met Glu Lys Thr Ser
 2820 2825 2830 2835

TCT GGA TTA TAC ATA TTT CGC AAT GAA AGA GAG GAA AAG GAA GCA 8781
 Ser Gly Leu Tyr Ile Phe Arg Asn Glu Arg Glu Glu Lys Glu Ala
 2840 2845 2850

GCA AAA TAT GTG GAG GCC CAA CAA AAG AGA CTA GAA GCC TTA TTC ACT 8829
 Ala Lys Tyr Val Glu Ala Gln Gln Lys Arg Leu Glu Ala Leu Phe Thr
 2855 2860 2865

AAA ATT CAG GAG GAA TTT GAA GAA CAT GAA GAA AAC ACA ACA AAA CCA 8877
 Lys Ile Gln Glu Phe Glu His Glu Glu Asn Thr Thr Lys Pro
 2870 2875 2880

TAT TTA CCA TCA CGT GCA CTA ACA AGA CAG CAA GTT CGT GCT TTG CAA 8925
 Tyr Leu Pro Ser Arg Ala Leu Thr Arg Gln Gln Val Arg Ala Leu Gln
 2885 2890 2895

GAT GGT GCA GAG CTT TAT GAA GCA GTG AAG AAT GCA GCA GAC CCA GCT 8973
 Asp Gly Ala Glu Leu Tyr Glu Ala Val Lys Asn Ala Ala Asp Pro Ala
 2900 2905 2910 2915

TAC CTT GAG GGT TAT TTC AGT GAA GAG CAG TTA AGA GCC TTG AAT AAT 9021
 Tyr Leu Glu Gly Tyr Phe Ser Glu Glu Gln Leu Arg Ala Leu Asn Asn
 2920 2925 2930

CAC AGG CAA ATG TTG AAT GAT AAG AAA CAA GCT CAG ATC CAG TTG GAA 9069
 His Arg Gln Met Leu Asn Asp Lys Lys Gln Ala Gln Ile Gln Leu Glu
 2935 2940 2945

ATT AGG AAG GCC ATG GAA TCT GCT GAA CAA AAG GAA CAA GGT TTA TCA 9117
 Ile Arg Lys Ala Met Glu Ser Ala Glu Gln Lys Glu Gln Gly Leu Ser
 2950 2955 2960

AGG GAT GTC ACA ACC GTG TGG AAG TTG CGT ATT GTA AGC TAT TCA AAA 9165
 Arg Asp Val Thr Thr Val Trp Lys Leu Arg Ile Val Ser Tyr Ser Lys
 2965 2970 2975

AAA GAA AAA GAT TCA GTT ATA CTG AGT ATT TGG CGT CCA TCA TCA GAT 9213
 Lys Glu Lys Asp Ser Val Ile Leu Ser Ile Trp Arg Pro Ser Ser Asp
 2980 2985 2990 2995

TTA TAT TCT CTG TTA ACA GAA GGA AAG AGA TAC AGA ATT TAT CAT CTT 9261
 Leu Tyr Ser Leu Leu Thr Glu Gly Lys Arg Tyr Arg Ile Tyr His Leu
 3000 3005 3010

GCA ACT TCA AAA TCT AAA AGT AAA TCT GAA AGA GCT AAC ATA CAG TTA 9309
 Ala Thr Ser Lys Ser Lys Ser Glu Arg Ala Asn Ile Gln Leu
 3015 3020 3025

GCA GCG ACA AAA AAA ACT CAG TAT CAA CAA CTA CCG GTT TCA GAT GAA 9357
 Ala Ala Thr Lys Lys Thr Gln Tyr Gln Gln Leu Pro Val Ser Asp Glu
 3030 3035 3040

ATT TTA TTT CAG ATT TAC CAG CCA CGG GAG CCC CTT CAC TTC AGC AAA 9405
 Ile Leu Phe Gln Ile Tyr Gln Pro Arg Glu Pro Leu His Phe Ser Lys
 3045 3050 3055

D
cont

| | |
|---|-------|
| TTT TTA GAT CCA GAC TTT CAG CCA TCT TGT GAG GTG GAC CTA ATA Phe Leu Asp Pro Asp Phe Gln Pro Ser Cys Ser Glu Val Asp Leu Ile 3060 3065 3070 3075 | 9453 |
| GGA TTT GTC GTT TCT GTG AAA AAA ACA GGA CTT GCC CCT TTC GTC Gly Phe Val Val Ser Val Val Lys Lys Thr Gly Leu Ala Pro Phe Val 3080 3085 3090 | 9501 |
| TAT TTG TCA GAC GAA TGT TAC AAT TTA CTG GCA ATA AAG TTT TGG ATA Tyr Leu Ser Asp Glu Cys Tyr Asn Leu Leu Ala Ile Lys Phe Trp Ile 3095 3100 3105 | 9549 |
| GAC CTT AAT GAG GAC ATT ATT AAG CCT CAT ATG TTA ATT GCT GCA AGC Asp Leu Asn Glu Asp Ile Ile Lys Pro His Met Leu Ile Ala Ala Ser 3110 3115 3120 | 9597 |
| AAC CTC CAG TGG CGA CCA GAA TCC AAA TCA GGC CTT CTT ACT TTA TTT Asn Leu Gln Trp Arg Pro Glu Ser Lys Ser Gly Leu Leu Thr Leu Phe 3125 3130 3135 | 9645 |
| GCT GGA GAT TTT TCT GTG TTT TCT GCT AGT CCA AAA GAG GGC CAC TTT Ala Gly Asp Phe Ser Val Phe Ser Ala Ser Pro Lys Glu Gly His Phe 3140 3145 3150 3155 | 9693 |
| CAA GAG ACA TTC AAC AAA ATG AAA AAT ACT GTT GAG AAT ATT GAC ATA Gln Glu Thr Phe Asn Lys Met Lys Asn Thr Val Glu Asn Ile Asp Ile 3160 3165 3170 | 9741 |
| CTT TGC AAT GAA GCA GAA AAC AAG CTT ATG CAT ATA CTG CAT GCA AAT Leu Cys Asn Glu Ala Glu Asn Lys Leu Met His Ile Leu His Ala Asn 3175 3180 3185 | 9789 |
| GAT CCC AAG TGG TCC ACC CCA ACT AAA GAC TGT ACT TCA GGG CCG TAC Asp Pro Lys Trp Ser Thr Pro Thr Lys Asp Cys Thr Ser Gly Pro Tyr 3190 3195 3200 | 9837 |
| ACT GCT CAA ATC ATT CCT GGT ACA GGA AAC AAG CTT CTG ATG TCT TCT Thr Ala Gln Ile Ile Pro Gly Thr Gly Asn Lys Leu Leu Met Ser Ser 3205 3210 3215 | 9885 |
| CCT AAT TGT GAG ATA TAT TAT CAA AGT CCT TTA TCA CTT TGT ATG GCC Pro Asn Cys Glu Ile Tyr Tyr Gln Ser Pro Leu Ser Leu Cys Met Ala 3220 3225 3230 3235 | 9933 |
| AAA AGG AAG TCT GTT TCC ACA CCT GTC TCA GCC CAG ATG ACT TCA AAG Lys Arg Lys Ser Val Ser Thr Pro Val Ser Ala Gln Met Thr Ser Lys 3240 3245 3250 | 9981 |
| TCT TGT AAA GGG GAG AAA GAG ATT GAT GAC CAA AAG AAC TGC AAA AAG Ser Cys Lys Gly Glu Lys Glu Ile Asp Asp Gln Lys Asn Cys Lys Lys 3255 3260 3265 | 10029 |
| AGA AGA GCC TTG GAT TTC TTG AGT AGA CTG CCT TTA CCT CCA CCT GTT Arg Arg Ala Leu Asp Phe Leu Ser Arg Leu Pro Leu Pro Pro Val 3270 3275 3280 | 10077 |
| AGT CCC ATT TGT ACA TTT GTT TCT CCG GCT GCA CAG AAG GCA TTT CAG Ser Pro Ile Cys Thr Phe Val Ser Pro Ala Ala Gln Lys Ala Phe Gln 3285 3290 3295 | 10125 |
| CCA CCA AGG AGT TGT GGC ACC AAA TAC GAA ACA CCC ATA AAG AAA AAA | 10173 |

Pro Pro Arg Ser Cys Gly Thr Lys Tyr Glu Thr Pro Ile Lys Lys Lys
 3300 3305 3310 3315
 GAA CTG AAT TCT CCT CAG ATG ACT CCA TTT AAA AAA TTC AAT GAA ATT 10221
 Glu Leu Asn Ser Pro Gln Met Thr Pro Phe Lys Lys Phe Asn Glu Ile
 3320 3325 3330
 TCT CTT TTG GAA AGT AAT TCA ATA GCT GAC GAA GAA CTT GCA TTG ATA 10269
 Ser Leu Leu Glu Ser Asn Ser Ile Ala Asp Glu Glu Leu Ala Leu Ile
 3335 3340 3345
 AAT ACC CAA GCT CTT TTG TCT GGT TCA ACA GGA GAA AAA CAA TTT ATA 10317
 Asn Thr Gln Ala Leu Leu Ser Gly Ser Thr Gly Glu Lys Gln Phe Ile
 3350 3355 3360
 TCT GTC AGT GAA TCC ACT AGG ACT GCT CCC ACC AGT TCA GAA GAT TAT 10365
 Ser Val Ser Glu Ser Thr Arg Thr Ala Pro Thr Ser Ser Glu Asp Tyr
 3365 3370 3375
 CTC AGA CTG AAA CGA CGT TGT ACT ACA TCT CTG ATC AAA GAA CAG GAG 10413
 Leu Arg Leu Lys Arg Arg Cys Thr Thr Ser Leu Ile Lys Glu Gln Glu
 3380 3385 3390 3395
 AGT TCC CAG GCC AGT ACG GAA GAA TGT GAG AAA AAT AAG CAG GAC ACA 10461
 Ser Ser Gln Ala Ser Thr Glu Glu Cys Glu Lys Asn Lys Gln Asp Thr
 3400 3405 3410
 ATT ACA ACT AAA AAA TAT ATC TAA 10485
 Ile Thr Thr Lys Lys Tyr Ile
 3415

D
cont

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 3418 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: protein
- (v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Met Pro Ile Gly Ser Lys Glu Arg Pro Thr Phe Phe Glu Ile Phe Lys
 1 5 10 15
 Thr Arg Cys Asn Lys Ala Asp Leu Gly Pro Ile Ser Leu Asn Trp Phe
 20 25 30
 Glu Glu Leu Ser Ser Glu Ala Pro Pro Tyr Asn Ser Glu Pro Ala Glu
 35 40 45
 Glu Ser Glu His Lys Asn Asn Tyr Glu Pro Asn Leu Phe Lys Thr
 50 55 60
 Pro Gln Arg Lys Pro Ser Tyr Asn Gln Leu Ala Ser Thr Pro Ile Ile
 65 70 75 80
 Phe Lys Glu Gln Gly Leu Thr Leu Pro Leu Tyr Gln Ser Pro Val Lys
 85 90 95
 Glu Leu Asp Lys Phe Lys Leu Asp Leu Gly Arg Asn Val Pro Asn Ser
 100 105 110
 Arg His Lys Ser Leu Arg Thr Val Lys Thr Lys Met Asp Gln Ala Asp
 115 120 125

Asp Val Ser Cys Pro Leu Leu Asn Ser Cys Leu Ser Glu Ser Pro Val
 130 135 140
 Val Leu Gln Cys Thr His Val Thr Pro Gln Arg Asp Lys Ser Val Val
 145 150 155 160
 Cys Gly Ser Leu Phe His Thr Pro Lys Phe Val Lys Gly Arg Gln Thr
 165 170 175
 Pro Lys His Ile Ser Glu Ser Leu Gly Ala Glu Val Asp Pro Asp Met
 180 185 190
 Ser Trp Ser Ser Ser Leu Ala Thr Pro Pro Thr Leu Ser Ser Thr Val
 195 200 205
 Leu Ile Val Arg Asn Glu Glu Ala Ser Glu Thr Val Phe Pro His Asp
 210 215 220
 Thr Thr Ala Asn Val Lys Ser Tyr Phe Ser Asn His Asp Glu Ser Leu
 225 230 235 240
 Lys Lys Asn Asp Arg Phe Ile Ala Ser Val Thr Asp Ser Glu Asn Thr
 245 250 255
 Asn Gln Arg Glu Ala Ala Ser His Gly Phe Gly Lys Thr Ser Gly Asn
 260 265 270
 Ser Phe Lys Val Asn Ser Cys Lys Asp His Ile Gly Lys Ser Met Pro
 275 280 285
 Asn Val Leu Glu Asp Glu Val Tyr Glu Thr Val Val Asp Thr Ser Glu
 290 295 300
 Glu Asp Ser Phe Ser Leu Cys Phe Ser Lys Cys Arg Thr Lys Asn Leu
 305 310 315 320
 Gln Lys Val Arg Thr Ser Lys Thr Arg Lys Lys Ile Phe His Glu Ala
 325 330 335
 Asn Ala Asp Glu Cys Glu Lys Ser Lys Asn Gln Val Lys Glu Lys Tyr
 340 345 350
 Ser Phe Val Ser Glu Val Glu Pro Asn Asp Thr Asp Pro Leu Asp Ser
 355 360 365
 Asn Val Ala His Gln Lys Pro Phe Glu Ser Gly Ser Asp Lys Ile Ser
 370 375 380
 Lys Glu Val Val Pro Ser Leu Ala Cys Glu Trp Ser Gln Leu Thr Leu
 385 390 395 400
 Ser Gly Leu Asn Gly Ala Gln Met Glu Lys Ile Pro Leu Leu His Ile
 405 410 415
 Ser Ser Cys Asp Gln Asn Ile Ser Glu Lys Asp Leu Leu Asp Thr Glu
 420 425 430
 Asn Lys Arg Lys Lys Asp Phe Leu Thr Ser Glu Asn Ser Leu Pro Arg
 435 440 445
 Ile Ser Ser Leu Pro Lys Ser Glu Lys Pro Leu Asn Glu Glu Thr Val
 450 455 460
 Val Asn Lys Arg Asp Glu Glu Gln His Leu Glu Ser His Thr Asp Cys
 465 470 475 480
 Ile Leu Ala Val Lys Gln Ala Ile Ser Gly Thr Ser Pro Val Ala Ser
 485 490 495
 Ser Phe Gln Gly Ile Lys Lys Ser Ile Phe Arg Ile Arg Glu Ser Pro
 500 505 510
 Lys Glu Thr Phe Asn Ala Ser Phe Ser Gly His Met Thr Asp Pro Asn
 515 520 525
 Phe Lys Lys Glu Thr Glu Ala Ser Glu Ser Gly Leu Glu Ile His Thr
 530 535 540
 Val Cys Ser Gln Lys Glu Asp Ser Leu Cys Pro Asn Leu Ile Asp Asn
 545 550 555 560
 Gly Ser Trp Pro Ala Thr Thr Gln Asn Ser Val Ala Leu Lys Asn
 565 570 575
 Ala Gly Leu Ile Ser Thr Leu Lys Lys Thr Asn Lys Phe Ile Tyr
 580 585 590
 Ala Ile His Asp Glu Thr Ser Tyr Lys Gly Lys Lys Ile Pro Lys Asp
 595 600 605
 Gln Lys Ser Glu Leu Ile Asn Cys Ser Ala Gln Phe Glu Ala Asn Ala

D
C
M

610 615 620
 Phe Glu Ala Pro Leu Thr Phe Ala Asn Ala Asp Ser Gly Leu Leu His
 625 630 635 640
 Ser Ser Val Lys Arg Ser Cys Ser Gln Asn Asp Ser Glu Glu Pro Thr
 645 650 655
 Leu Ser Leu Thr Ser Ser Phe Gly Thr Ile Leu Arg Lys Cys Ser Arg
 660 665 670
 Asn Glu Thr Cys Ser Asn Asn Thr Val Ile Ser Gln Asp Leu Asp Tyr
 675 680 685
 Lys Glu Ala Lys Cys Asn Lys Glu Lys Leu Gln Phe Ile Thr Pro
 690 695 700
 Glu Ala Asp Ser Leu Ser Cys Leu Gln Glu Gly Gln Cys Glu Asn Asp
 705 710 715 720
 Pro Lys Ser Lys Lys Val Ser Asp Ile Lys Glu Glu Val Leu Ala Ala
 725 730 735
 Ala Cys His Pro Val Gln His Ser Lys Val Glu Tyr Ser Asp Thr Asp
 740 745 750
 Phe Gln Ser Gln Lys Ser Leu Leu Tyr Asp His Glu Asn Ala Ser Thr
 755 760 765
 Leu Ile Leu Thr Pro Thr Ser Lys Asp Val Leu Ser Asn Leu Val Met
 770 775 780
 Ile Ser Arg Gly Lys Glu Ser Tyr Lys Met Ser Asp Lys Leu Lys Gly
 785 790 795 800
 Asn Asn Tyr Glu Ser Asp Val Glu Leu Thr Lys Asn Ile Pro Met Glu
 805 810 815
 Lys Asn Gln Asp Val Cys Ala Leu Asn Glu Asn Tyr Lys Asn Val Glu
 820 825 830
 Leu Leu Pro Pro Glu Lys Tyr Met Arg Val Ala Ser Pro Ser Arg Lys
 835 840 845
 Val Gln Phe Asn Gln Asn Thr Asn Leu Arg Val Ile Gln Lys Asn Gln
 850 855 860
 Glu Glu Thr Thr Ser Ile Ser Lys Ile Thr Val Asn Pro Asp Ser Glu
 865 870 875 880
 Glu Leu Phe Ser Asp Asn Glu Asn Asn Phe Val Phe Gln Val Ala Asn
 885 890 895
 Glu Arg Asn Asn Leu Ala Leu Gly Asn Thr Lys Glu Leu His Glu Thr
 900 905 910
 Asp Leu Thr Cys Val Asn Glu Pro Ile Phe Lys Asn Ser Thr Met Val
 915 920 925
 Leu Tyr Gly Asp Thr Gly Asp Lys Gln Ala Thr Gln Val Ser Ile Lys
 930 935 940
 Lys Asp Leu Val Tyr Val Leu Ala Glu Glu Asn Lys Asn Ser Val Lys
 945 950 955 960
 Gln His Ile Lys Met Thr Leu Gly Gln Asp Leu Lys Ser Asp Ile Ser
 965 970 975
 Leu Asn Ile Asp Lys Ile Pro Glu Lys Asn Asn Asp Tyr Met Asn Lys
 980 985 990
 Trp Ala Gly Leu Leu Gly Pro Ile Ser Asn His Ser Phe Gly Gly Ser
 995 1000 1005
 Phe Arg Thr Ala Ser Asn Lys Glu Ile Lys Leu Ser Glu His Asn Ile
 1010 1015 1020
 Lys Lys Ser Lys Met Phe Phe Lys Asp Ile Glu Glu Gln Tyr Pro Thr
 1025 1030 1035 1040
 Ser Leu Ala Cys Val Glu Ile Val Asn Thr Leu Ala Leu Asp Asn Gln
 1045 1050 1055
 Lys Lys Leu Ser Lys Pro Gln Ser Ile Asn Thr Val Ser Ala His Leu
 1060 1065 1070
 Gln Ser Ser Val Val Val Ser Asp Cys Lys Asn Ser His Ile Thr Pro
 1075 1080 1085
 Gln Met Leu Phe Ser Lys Gln Asp Phe Asn Ser Asn His Asn Leu Thr
 1090 1095 1100

Pro Ser Gln Lys Ala Glu Ile Thr Glu Leu Ser Thr Ile Leu Glu Glu
1105 1110 1115 1120
Ser Gly Ser Gln Phe Glu Phe Thr Gln Phe Arg Lys Pro Ser Tyr Ile
1125 1130 1135
Leu Gln Lys Ser Thr Phe Glu Val Pro Glu Asn Gln Met Thr Ile Leu
1140 1145 1150
Lys Thr Thr Ser Glu Glu Cys Arg Asp Ala Asp Leu His Val Ile Met
1155 1160 1165
Asn Ala Pro Ser Ile Gly Gln Val Asp Ser Ser Lys Gln Phe Glu Gly
1170 1175 1180
Thr Val Glu Ile Lys Arg Lys Phe Ala Gly Leu Leu Lys Asn Asp Cys
1185 1190 1195 1200
Asn Lys Ser Ala Ser Gly Tyr Leu Thr Asp Glu Asn Glu Val Gly Phe
1205 1210 1215
Arg Gly Phe Tyr Ser Ala His Gly Thr Lys Leu Asn Val Ser Thr Glu
1220 1225 1230
Ala Leu Gln Lys Ala Val Lys Leu Phe Ser Asp Ile Glu Asn Ile Ser
1235 1240 1245
Glu Glu Thr Ser Ala Glu Val His Pro Ile Ser Leu Ser Ser Ser Lys
1250 1255 1260
Cys His Asp Ser Val Val Ser Met Phe Lys Ile Glu Asn His Asn Asp
1265 1270 1275 1280
Lys Thr Val Ser Glu Lys Asn Asn Lys Cys Gln Leu Ile Leu Gln Asn
1285 1290 1295
Asn Ile Glu Met Thr Thr Gly Thr Phe Val Glu Glu Ile Thr Glu Asn
1300 1305 1310
Tyr Lys Arg Asn Thr Glu Asn Glu Asp Asn Lys Tyr Thr Ala Ala Ser
1315 1320 1325
Arg Asn Ser His Asn Leu Glu Phe Asp Gly Ser Asp Ser Ser Lys Asn
1330 1335 1340
Asp Thr Val Cys Ile His Lys Asp Glu Thr Asp Leu Leu Phe Thr Asp
1345 1350 1355 1360
Gln His Asn Ile Cys Leu Lys Leu Ser Gly Gln Phe Met Lys Glu Gly
1365 1370 1375
Asn Thr Gln Ile Lys Glu Asp Leu Ser Asp Leu Thr Phe Leu Glu Val
1380 1385 1390
Ala Lys Ala Gln Glu Ala Cys His Gly Asn Thr Ser Asn Lys Glu Gln
1395 1400 1405
Leu Thr Ala Thr Lys Thr Glu Gln Asn Ile Lys Asp Phe Glu Thr Ser
1410 1415 1420
Asp Thr Phe Phe Gln Thr Ala Ser Gly Lys Asn Ile Ser Val Ala Lys
1425 1430 1435 1440
Glu Ser Phe Asn Lys Ile Val Asn Phe Phe Asp Gln Lys Pro Glu Glu
1445 1450 1455
Leu His Asn Phe Ser Leu Asn Ser Glu Leu His Ser Asp Ile Arg Lys
1460 1465 1470
Asn Lys Met Asp Ile Leu Ser Tyr Glu Glu Thr Asp Ile Val Lys His
1475 1480 1485
Lys Ile Leu Lys Glu Ser Val Pro Val Gly Thr Gly Asn Gln Leu Val
1490 1495 1500
Thr Phe Gln Gly Gln Pro Glu Arg Asp Glu Lys Ile Lys Glu Pro Thr
1505 1510 1515 1520
Leu Leu Gly Phe His Thr Ala Ser Gly Lys Lys Val Lys Ile Ala Lys
1525 1530 1535
Glu Ser Leu Asp Lys Val Lys Asn Leu Phe Asp Glu Lys Glu Gln Gly
1540 1545 1550
Thr Ser Glu Ile Thr Ser Phe Ser His Gln Trp Ala Lys Thr Leu Lys
1555 1560 1565
Tyr Arg Glu Ala Cys Lys Asp Leu Glu Leu Ala Cys Glu Thr Ile Glu
1570 1575 1580
Ile Thr Ala Ala Pro Lys Cys Lys Glu Met Gln Asn Ser Leu Asn Asn

1585 1590 1595 1600
 Asp Lys Asn Leu Val Ser Ile Glu Thr Val Val Pro Pro Lys Leu Leu
 1605 1610 1615
 Ser Asp Asn Leu Cys Arg Gln Thr Glu Asn Leu Lys Thr Ser Lys Ser
 1620 1625 1630
 Ile Phe Leu Lys Val Lys Val His Glu Asn Val Glu Lys Glu Thr Ala
 1635 1640 1645
 Lys Ser Pro Ala Thr Cys Tyr Thr Asn Gln Ser Pro Tyr Ser Val Ile
 1650 1655 1660
 Glu Asn Ser Ala Leu Ala Phe Tyr Thr Ser Cys Ser Arg Lys Thr Ser
 1665 1670 1675 1680
 Val Ser Gln Thr Ser Leu Leu Glu Ala Lys Lys Trp Leu Arg Glu Gly
 1685 1690 1695
 Ile Phe Asp Gly Gln Pro Glu Arg Ile Asn Thr Ala Asp Tyr Val Gly
 1700 1705 1710
 Asn Tyr Leu Tyr Glu Asn Asn Ser Asn Ser Thr Ile Ala Glu Asn Asp
 1715 1720 1725
 Lys Asn His Leu Ser Glu Lys Gln Asp Thr Tyr Leu Ser Asn Ser Ser
 1730 1735 1740
 Met Ser Asn Ser Tyr Ser Tyr His Ser Asp Glu Val Tyr Asn Asp Ser
 1745 1750 1755 1760
 Gly Tyr Leu Ser Lys Asn Lys Leu Asp Ser Gly Ile Glu Pro Val Leu
 1765 1770 1775
 Lys Asn Val Glu Asp Gln Lys Asn Thr Ser Phe Ser Lys Val Ile Ser
 1780 1785 1790
 Asn Val Lys Asp Ala Asn Ala Tyr Pro Gln Thr Val Asn Glu Asp Ile
 1795 1800 1805
 Cys Val Glu Glu Leu Val Thr Ser Ser Pro Cys Lys Asn Lys Asn
 1810 1815 1820
 Ala Ala Ile Lys Leu Ser Ile Ser Asn Ser Asn Asn Phe Glu Val Gly
 1825 1830 1835 1840
 Pro Pro Ala Phe Arg Ile Ala Ser Gly Lys Ile Val Cys Val Ser His
 1845 1850 1855
 Glu Thr Ile Lys Lys Val Lys Asp Ile Phe Thr Asp Ser Phe Ser Lys
 1860 1865 1870
 Val Ile Lys Glu Asn Asn Glu Asn Lys Ser Lys Ile Cys Gln Thr Lys
 1875 1880 1885
 Ile Met Ala Gly Cys Tyr Glu Ala Leu Asp Asp Ser Glu Asp Ile Leu
 1890 1895 1900
 His Asn Ser Leu Asp Asn Asp Glu Cys Ser Thr His Ser His Lys Val
 1905 1910 1915 1920
 Phe Ala Asp Ile Gln Ser Glu Glu Ile Leu Gln His Asn Gln Asn Met
 1925 1930 1935
 Ser Gly Leu Glu Lys Val Ser Lys Ile Ser Pro Cys Asp Val Ser Leu
 1940 1945 1950
 Glu Thr Ser Asp Ile Cys Lys Cys Ser Ile Gly Lys Leu His Lys Ser
 1955 1960 1965
 Val Ser Ser Ala Asn Thr Cys Gly Ile Phe Ser Thr Ala Ser Gly Lys
 1970 1975 1980
 Ser Val Gln Val Ser Asp Ala Ser Leu Gln Asn Ala Arg Gln Val Phe
 1985 1990 1995 2000
 Ser Glu Ile Glu Asp Ser Thr Lys Gln Val Phe Ser Lys Val Leu Phe
 2005 2010 2015
 Lys Ser Asn Glu His Ser Asp Gln Leu Thr Arg Glu Glu Asn Thr Ala
 2020 2025 2030
 Ile Arg Thr Pro Glu His Leu Ile Ser Gln Lys Gly Phe Ser Tyr Asn
 2035 2040 2045
 Val Val Asn Ser Ser Ala Phe Ser Gly Phe Ser Thr Ala Ser Gly Lys
 2050 2055 2060
 Gln Val Ser Ile Leu Glu Ser Ser Leu His Lys Val Lys Gly Val Leu
 2065 2070 2075 2080

Glu Glu Phe Asp Leu Ile Arg Thr Glu His Ser Leu His Tyr Ser Pro
 2085 2090 2095
 Thr Ser Arg Gln Asn Val Ser Lys Ile Leu Pro Arg Val Asp Lys Arg
 2100 2105 2110
 Asn Pro Glu His Cys Val Asn Ser Glu Met Glu Lys Thr Cys Ser Lys
 2115 2120 2125
 Glu Phe Lys Leu Ser Asn Asn Leu Asn Val Glu Gly Gly Ser Ser Glu
 2130 2135 2140
 Asn Asn His Ser Ile Lys Val Ser Pro Tyr Leu Ser Gln Phe Gln Gln
 2145 2150 2155 2160
 Asp Lys Gln Gln Leu Val Leu Gly Thr Lys Val Ser Leu Val Glu Asn
 2165 2170 2175
 Ile His Val Leu Gly Lys Glu Gln Ala Ser Pro Lys Asn Val Lys Met
 2180 2185 2190
 Glu Ile Gly Lys Thr Glu Thr Phe Ser Asp Val Pro Val Lys Thr Asn
 2195 2200 2205
 Ile Glu Val Cys Ser Thr Tyr Ser Lys Asp Ser Glu Asn Tyr Phe Glu
 2210 2215 2220
 Thr Glu Ala Val Glu Ile Ala Lys Ala Phe Met Glu Asp Asp Glu Leu
 2225 2230 2235 2240
 Thr Asp Ser Lys Leu Pro Ser His Ala Thr His Ser Leu Phe Thr Cys
 2245 2250 2255
 Pro Glu Asn Glu Glu Met Val Leu Ser Asn Ser Arg Ile Gly Lys Arg
 2260 2265 2270
 Arg Gly Glu Pro Leu Ile Leu Val Gly Glu Pro Ser Ile Lys Arg Asn
 2275 2280 2285
 Leu Leu Asn Glu Phe Asp Arg Ile Ile Glu Asn Gln Glu Lys Ser Leu
 2290 2295 2300
 Lys Ala Ser Lys Ser Thr Pro Asp Gly Thr Ile Lys Asp Arg Arg Leu
 2305 2310 2315 2320
 Phe Met His His Val Ser Leu Glu Pro Ile Thr Cys Val Pro Phe Arg
 2325 2330 2335
 Thr Thr Lys Glu Arg Gln Glu Ile Gln Asn Pro Asn Phe Thr Ala Pro
 2340 2345 2350
 Gly Gln Glu Phe Leu Ser Lys Ser His Leu Tyr Glu His Leu Thr Leu
 2355 2360 2365
 Glu Lys Ser Ser Ser Asn Leu Ala Val Ser Gly His Pro Phe Tyr Gln
 2370 2375 2380
 Val Ser Ala Thr Arg Asn Glu Lys Met Arg His Leu Ile Thr Thr Gly
 2385 2390 2395 2400
 Arg Pro Thr Lys Val Phe Val Pro Pro Phe Lys Thr Lys Ser His Phe
 2405 2410 2415
 His Arg Val Glu Gln Cys Val Arg Asn Ile Asn Leu Glu Glu Asn Arg
 2420 2425 2430
 Gln Lys Gln Asn Ile Asp Gly His Gly Ser Asp Asp Ser Lys Asn Lys
 2435 2440 2445
 Ile Asn Asp Asn Glu Ile His Gln Phe Asn Lys Asn Asn Ser Asn Gln
 2450 2455 2460
 Ala Ala Ala Val Thr Phe Thr Lys Cys Glu Glu Pro Leu Asp Leu
 2465 2470 2475 2480
 Ile Thr Ser Leu Gln Asn Ala Arg Asp Ile Gln Asp Met Arg Ile Lys
 2485 2490 2495
 Lys Lys Gln Arg Gln Arg Val Phe Pro Gln Pro Gly Ser Leu Tyr Leu
 2500 2505 2510
 Ala Lys Thr Ser Thr Leu Pro Arg Ile Ser Leu Lys Ala Ala Val Gly
 2515 2520 2525
 Gly Gln Val Pro Ser Ala Cys Ser His Lys Gln Leu Tyr Thr Tyr Gly
 2530 2535 2540
 Val Ser Lys His Cys Ile Lys Ile Asn Ser Lys Asn Ala Glu Ser Phe
 2545 2550 2555 2560
 Gln Phe His Thr Glu Asp Tyr Phe Gly Lys Glu Ser Leu Trp Thr Gly

2565 2570 2575
 Lys Gly Ile Gln Leu Ala Asp Gly Gly Trp Leu Ile Pro Ser Asn Asp
 2580 2585 2590
 Gly Lys Ala Gly Lys Glu Glu Phe Tyr Arg Ala Leu Cys Asp Thr Pro
 2595 2600 2605
 Gly Val Asp Pro Lys Leu Ile Ser Arg Ile Trp Val Tyr Asn His Tyr
 2610 2615 2620
 Arg Trp Ile Ile Trp Lys Leu Ala Ala Met Glu Cys Ala Phe Pro Lys
 2625 2630 2635 2640
 Glu Phe Ala Asn Arg Cys Leu Ser Pro Glu Arg Val Leu Leu Gln Leu
 2645 2650 2655
 Lys Tyr Arg Tyr Asp Thr Glu Ile Asp Arg Ser Arg Arg Ser Ala Ile
 2660 2665 2670
 Lys Lys Ile Met Glu Arg Asp Asp Thr Ala Ala Lys Thr Leu Val Leu
 2675 2680 2685
 Cys Val Ser Asp Ile Ile Ser Leu Ser Ala Asn Ile Ser Glu Thr Ser
 2690 2695 2700
 Ser Asn Lys Thr Ser Ala Asp Thr Gln Lys Val Ala Ile Ile Glu
 2705 2710 2715 2720
 Leu Thr Asp Gly Trp Tyr Ala Val Lys Ala Gln Leu Asp Pro Pro Leu
 2725 2730 2735
 Leu Ala Val Leu Lys Asn Gly Arg Leu Thr Val Gly Gln Lys Ile Ile
 2740 2745 2750
 Leu His Gly Ala Glu Leu Val Gly Ser Pro Asp Ala Cys Thr Pro Leu
 2755 2760 2765
 Glu Ala Pro Glu Ser Leu Met Leu Lys Ile Ser Ala Asn Ser Thr Arg
 2770 2775 2780
 Pro Ala Arg Trp Tyr Thr Lys Leu Gly Phe Phe Pro Asp Pro Arg Pro
 2785 2790 2795 2800
 Phe Pro Leu Pro Leu Ser Ser Leu Phe Ser Asp Gly Gly Asn Val Gly
 2805 2810 2815
 Cys Val Asp Val Ile Ile Gln Arg Ala Tyr Pro Ile Gln Trp Met Glu
 2820 2825 2830
 Lys Thr Ser Ser Gly Leu Tyr Ile Phe Arg Asn Glu Arg Glu Glu Glu
 2835 2840 2845
 Lys Glu Ala Ala Lys Tyr Val Glu Ala Gln Gln Lys Arg Leu Glu Ala
 2850 2855 2860
 Leu Phe Thr Lys Ile Gln Glu Glu Phe Glu Glu His Glu Glu Asn Thr
 2865 2870 2875 2880
 Thr Lys Pro Tyr Leu Pro Ser Arg Ala Leu Thr Arg Gln Gln Val Arg
 2885 2890 2895
 Ala Leu Gln Asp Gly Ala Glu Leu Tyr Glu Ala Val Lys Asn Ala Ala
 2900 2905 2910
 Asp Pro Ala Tyr Leu Glu Gly Tyr Phe Ser Glu Glu Gln Leu Arg Ala
 2915 2920 2925
 Leu Asn Asn His Arg Gln Met Leu Asn Asp Lys Lys Gln Ala Gln Ile
 2930 2935 2940
 Gln Leu Glu Ile Arg Lys Ala Met Glu Ser Ala Glu Gln Lys Glu Gln
 2945 2950 2955 2960
 Gly Leu Ser Arg Asp Val Thr Thr Val Trp Lys Leu Arg Ile Val Ser
 2965 2970 2975
 Tyr Ser Lys Lys Glu Lys Asp Ser Val Ile Leu Ser Ile Trp Arg Pro
 2980 2985 2990
 Ser Ser Asp Leu Tyr Ser Leu Leu Thr Glu Gly Lys Arg Tyr Arg Ile
 2995 3000 3005
 Tyr His Leu Ala Thr Ser Lys Ser Lys Ser Glu Arg Ala Asn
 3010 3015 3020
 Ile Gln Leu Ala Ala Thr Lys Lys Thr Gln Tyr Gln Gln Leu Pro Val
 3025 3030 3035 3040
 Ser Asp Glu Ile Leu Phe Gln Ile Tyr Gln Pro Arg Glu Pro Leu His
 3045 3050 3055

Phe Ser Lys Phe Leu Asp Pro Asp Phe Gln Pro Ser Cys Ser Glu Val
 3060 3065 3070
 Asp Leu Ile Gly Phe Val Val Ser Val Val Lys Lys Thr Gly Leu Ala
 3075 3080 3085
 Pro Phe Val Tyr Leu Ser Asp Glu Cys Tyr Asn Leu Leu Ala Ile Lys
 3090 3095 3100
 Phe Trp Ile Asp Leu Asn Glu Asp Ile Ile Lys Pro His Met Leu Ile
 3105 3110 3115 3120
 Ala Ala Ser Asn Leu Gln Trp Arg Pro Glu Ser Lys Ser Gly Leu Leu
 3125 3130 3135
 Thr Leu Phe Ala Gly Asp Phe Ser Val Phe Ser Ala Ser Pro Lys Glu
 3140 3145 3150
 Gly His Phe Gln Glu Thr Phe Asn Lys Met Lys Asn Thr Val Glu Asn
 3155 3160 3165
 Ile Asp Ile Leu Cys Asn Glu Ala Glu Asn Lys Leu Met His Ile Leu
 3170 3175 3180
 His Ala Asn Asp Pro Lys Trp Ser Thr Pro Thr Lys Asp Cys Thr Ser
 3185 3190 3195 3200
 Gly Pro Tyr Thr Ala Gln Ile Ile Pro Gly Thr Gly Asn Lys Leu Leu
 3205 3210 3215
 Met Ser Ser Pro Asn Cys Glu Ile Tyr Tyr Gln Ser Pro Leu Ser Leu
 3220 3225 3230
 Cys Met Ala Lys Arg Lys Ser Val Ser Thr Pro Val Ser Ala Gln Met
 3235 3240 3245
 Thr Ser Lys Ser Cys Lys Gly Glu Lys Glu Ile Asp Asp Gln Lys Asn
 3250 3255 3260
 Cys Lys Lys Arg Arg Ala Leu Asp Phe Leu Ser Arg Leu Pro Leu Pro
 3265 3270 3275 3280
 Pro Pro Val Ser Pro Ile Cys Thr Phe Val Ser Pro Ala Ala Gln Lys
 3285 3290 3295
 Ala Phe Gln Pro Pro Arg Ser Cys Gly Thr Lys Tyr Glu Thr Pro Ile
 3300 3305 3310
 Lys Lys Lys Glu Leu Asn Ser Pro Gln Met Thr Pro Phe Lys Lys Phe
 3315 3320 3325
 Asn Glu Ile Ser Leu Leu Glu Ser Asn Ser Ile Ala Asp Glu Glu Leu
 3330 3335 3340
 Ala Leu Ile Asn Thr Gln Ala Leu Leu Ser Gly Ser Thr Gly Glu Lys
 3345 3350 3355 3360
 Gln Phe Ile Ser Val Ser Glu Ser Thr Arg Thr Ala Pro Thr Ser Ser
 3365 3370 3375
 Glu Asp Tyr Leu Arg Leu Lys Arg Arg Cys Thr Thr Ser Leu Ile Lys
 3380 3385 3390
 Glu Gln Glu Ser Ser Gln Ala Ser Thr Glu Glu Cys Glu Lys Asn Lys
 3395 3400 3405
 Gln Asp Thr Ile Thr Thr Lys Lys Tyr Ile
 3410 3415

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 10485 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (ix) FEATURE:
 - (A) NAME/KEY: Coding Sequence
 - (B) LOCATION: 229...10482
 - (D) OTHER INFORMATION: BRCA2 (OMI2)

(ix) FEATURE:

- (A) NAME/KEY: variation
 (B) LOCATION: 3624
 (D) OTHER INFORMATION: R = A or G. Xaa (amino acid position 1132) = Lys.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GGTGGCGCGA CCTTCTGAAA CTAGGCGGCA GAGGCGGAGC CGCTGTGGCA CTGCTGCGCC 60
 TCTGCTGCGC CTCGGGTGTC TTTGCGGCG GTGGGTCGCC GCCGGGAGAA GCGTGAGGGG 120
 ACAGATTGT GACCGGGCGCG GTTTTGTCA GCTTACTCCG GCCAAAAAAAG AACTGCACCT 180
 CTGGAGCGGA CTTATTTACC AAGCATTGGA GGAATATCGT AGGTAAAA ATG CCT ATT 237
 Met Pro Ile
 1

GGA TCC AAA GAG AGG CCA ACA TTT TTT GAA ATT TTT AAG ACA CGC TGC 285
 Gly Ser Lys Glu Arg Pro Thr Phe Phe Glu Ile Phe Lys Thr Arg Cys
 5 10 15

AAC AAA GCA GAT TTA GGA CCA ATA AGT CTT AAT TGG TTT GAA GAA CTT 333
 Asn Lys Ala Asp Leu Gly Pro Ile Ser Leu Asn Trp Phe Glu Glu Leu
 20 25 30 35

TCT TCA GAA GCT CCA CCC TAT AAT TCT GAA CCT GCA GAA GAA TCT GAA 381
 Ser Ser Glu Ala Pro Pro Tyr Asn Ser Glu Pro Ala Glu Glu Ser Glu
 40 45 50

CAT AAA AAC AAC AAT TAC GAA CCA AAC CTA TTT AAA ACT CCA CAA AGG 429
 His Lys Asn Asn Asn Tyr Glu Pro Asn Leu Phe Lys Thr Pro Gln Arg
 55 60 65

AAA CCA TCT TAT AAT CAG CTG GCT TCA ACT CCA ATA ATA TTC AAA GAG 477
 Lys Pro Ser Tyr Asn Gln Leu Ala Ser Thr Pro Ile Ile Phe Lys Glu
 70 75 80

CAA GGG CTG ACT CTG CCG CTG TAC CAA TCT CCT GTA AAA GAA TTA GAT 525
 Gln Gly Leu Thr Leu Pro Leu Tyr Gln Ser Pro Val Lys Glu Leu Asp
 85 90 95

AAA TTC AAA TTA GAC TTA GGA AGG AAT GTT CCC AAT AGT AGA CAT AAA 573
 Lys Phe Lys Leu Asp Leu Gly Arg Asn Val Pro Asn Ser Arg His Lys
 100 105 110 115

AGT CTT CGC ACA GTG AAA ACT AAA ATG GAT CAA GCA GAT GAT GTT TCC 621
 Ser Leu Arg Thr Val Lys Thr Lys Met Asp Gln Ala Asp Asp Val Ser
 120 125 130

TGT CCA CTT CTA AAT TCT TGT CTT AGT GAA AGT CCT GTT GTT CTA CAA 669
 Cys Pro Leu Leu Asn Ser Cys Leu Ser Glu Ser Pro Val Val Leu Gln
 135 140 145

TGT ACA CAT GTA ACA CCA CAA AGA GAT AAG TCA GTG GTA TGT GGG AGT 717
 Cys Thr His Val Thr Pro Gln Arg Asp Lys Ser Val Val Cys Gly Ser
 150 155 160

TTG TTT CAT ACA CCA AAG TTT GTG AAG GGT CGT CAG ACA CCA AAA CAT 765
 Leu Phe His Thr Pro Lys Phe Val Lys Gly Arg Gln Thr Pro Lys His
 165 170 175

ATT TCT GAA AGT CTA GGA GCT GAG GTG GAT CCT GAT ATG TCT TGG TCA 813
 Ile Ser Glu Ser Leu Gly Ala Glu Val Asp Pro Asp Met Ser Trp Ser

180 185 190 195

| | |
|--|------|
| AGT TCT TTA GCT ACA CCA CCC ACC CTT AGT TCT ACT GTG CTC ATA GTC Ser Ser Leu Ala Thr Pro Pro Thr Leu Ser Ser Thr Val Leu Ile Val 200 205 210 | 861 |
| AGA AAT GAA GAA GCA TCT GAA ACT GTA TTT CCT CAT GAT ACT ACT GCT Arg Asn Glu Glu Ala Ser Glu Thr Val Phe Pro His Asp Thr Thr Ala 215 220 225 | 909 |
| AAT GTG AAA AGC TAT TTT TCC AAT CAT GAT GAA AGT CTG AAG AAA AAT Asn Val Lys Ser Tyr Phe Ser Asn His Asp Glu Ser Leu Lys Lys Asn 230 235 240 | 957 |
| GAT AGA TTT ATC GCT TCT GTG ACA GAC AGT GAA AAC ACA AAT CAA AGA Asp Arg Phe Ile Ala Ser Val Thr Asp Ser Glu Asn Thr Asn Gln Arg 245 250 255 | 1005 |
| GAA GCT GCA AGT CAT GGA TTT GGA AAA ACA TCA GGG AAT TCA TTT AAA Glu Ala Ala Ser His Gly Phe Gly Lys Thr Ser Gly Asn Ser Phe Lys 260 265 270 275 | 1053 |
| GTA AAT AGC TGC AAA GAC CAC ATT GGA AAG TCA ATG CCA AAT GTC CTA Val Asn Ser Cys Lys Asp His Ile Gly Lys Ser Met Pro Asn Val Leu 280 285 290 | 1101 |
| GAA GAT GAA GTA TAT GAA ACA GTT GTA GAT ACC TCT GAA GAA GAT AGT Glu Asp Glu Val Tyr Glu Thr Val Val Asp Thr Ser Glu Glu Asp Ser 295 300 305 | 1149 |
| TTT TCA TTA TGT TTT TCT AAA TGT AGA ACA AAA AAT CTA CAA AAA GTA Phe Ser Leu Cys Phe Ser Lys Cys Arg Thr Lys Asn Leu Gln Lys Val 310 315 320 | 1197 |
| AGA ACT AGC AAG ACT AGG AAA AAA ATT TTC CAT GAA GCA AAC GCT GAT Arg Thr Ser Lys Thr Arg Lys Lys Ile Phe His Glu Ala Asn Ala Asp 325 330 335 | 1245 |
| GAA TGT GAA AAA TCT AAA AAC CAA GTG AAA GAA AAA TAC TCA TTT GTA Glu Cys Glu Lys Ser Lys Asn Gln Val Lys Glu Lys Tyr Ser Phe Val 340 345 350 355 | 1293 |
| TCT GAA GTG GAA CCA AAT GAT ACT GAT CCA TTA GAT TCA AAT GTA GCA Ser Glu Val Glu Pro Asn Asp Thr Asp Pro Leu Asp Ser Asn Val Ala 360 365 370 | 1341 |
| CAT CAG AAG CCC TTT GAG AGT GGA AGT GAC AAA ATC TCC AAG GAA GTT His Gln Lys Pro Phe Glu Ser Gly Ser Asp Lys Ile Ser Lys Glu Val 375 380 385 | 1389 |
| GTA CCG TCT TTG GCC TGT GAA TGG TCT CAA CTA ACC CTT TCA GGT CTA Val Pro Ser Leu Ala Cys Glu Trp Ser Gln Leu Thr Leu Ser Gly Leu 390 395 400 | 1437 |
| AAT GGA GCC CAG ATG GAG AAA ATA CCC CTA TTG CAT ATT TCT TCA TGT Asn Gly Ala Gln Met Glu Lys Ile Pro Leu Leu His Ile Ser Ser Cys 405 410 415 | 1485 |
| GAC CAA AAT ATT TCA GAA AAA GAC CTA TTA GAC ACA GAG AAC AAA AGA Asp Gln Asn Ile Ser Glu Lys Asp Leu Leu Asp Thr Glu Asn Lys Arg 420 425 430 435 | 1533 |

AAG AAA GAT TTT CTT ACT TCA GAG AAT TCT TTG CCA CGT ATT TCT AGC 1581
 Lys Lys Asp Phe Leu Thr Ser Glu Asn Ser Leu Pro Arg Ile Ser Ser
 440 445 450

 CTA CCA AAA TCA GAG AAG CCA TTA AAT GAG GAA ACA GTG GTA AAT AAG 1629
 Leu Pro Lys Ser Glu Lys Pro Leu Asn Glu Glu Thr Val Val Asn Lys
 455 460 465

 AGA GAT GAA GAG CAG CAT CTT GAA TCT CAT ACA GAC TGC ATT CTT GCA 1677
 Arg Asp Glu Glu Gln His Leu Glu Ser His Thr Asp Cys Ile Leu Ala
 470 475 480

 GTA AAG CAG GCA ATA TCT GGA ACT TCT CCA GTG GCT TCT TCA TTT CAG 1725
 Val Lys Gln Ala Ile Ser Gly Thr Ser Pro Val Ala Ser Ser Phe Gln
 485 490 495

 GGT ATC AAA AAG TCT ATA TTC AGA ATA AGA GAA TCA CCT AAA GAG ACT 1773
 Gly Ile Lys Lys Ser Ile Phe Arg Ile Arg Glu Ser Pro Lys Glu Thr
 500 505 510 515

 TTC AAT GCA AGT TTT TCA GGT CAT ATG ACT GAT CCA AAC TTT AAA AAA 1821
 Phe Asn Ala Ser Phe Ser Gly His Met Thr Asp Pro Asn Phe Lys Lys
 520 525 530


 GAA ACT GAA GCC TCT GAA AGT GGA CTG GAA ATA CAT ACT GTT TGC TCA 1869
 Glu Thr Glu Ala Ser Glu Ser Gly Leu Glu Ile His Thr Val Cys Ser
 535 540 545

 CAG AAG GAG GAC TCC TTA TGT CCA AAT TTA ATT GAT AAT GGA AGC TGG 1917
 Gln Lys Glu Asp Ser Leu Cys Pro Asn Leu Ile Asp Asn Gly Ser Trp
 550 555 560

 CCA GCC ACC ACC ACA CAG AAT TCT GTA GCT TTG AAG AAT GCA GGT TTA 1965
 Pro Ala Thr Thr Gln Asn Ser Val Ala Leu Lys Asn Ala Gly Leu
 565 570 575

 ATA TCC ACT TTG AAA AAG AAA ACA AAT AAG TTT ATT TAT GCT ATA CAT 2013
 Ile Ser Thr Leu Lys Lys Thr Asn Lys Phe Ile Tyr Ala Ile His
 580 585 590 595

 GAT GAA ACA TCT TAT AAA GGA AAA AAA ATA CCG AAA GAC CAA AAA TCA 2061
 Asp Glu Thr Ser Tyr Lys Gly Lys Lys Ile Pro Lys Asp Gln Lys Ser
 600 605 610

 GAA CTA ATT AAC TGT TCA GCC CAG TTT GAA GCA AAT GCT TTT GAA GCA 2109
 Glu Leu Ile Asn Cys Ser Ala Gln Phe Glu Ala Asn Ala Phe Glu Ala
 615 620 625

 CCA CTT ACA TTT GCA AAT GCT GAT TCA GGT TTA TTG CAT TCT TCT GTG 2157
 Pro Leu Thr Phe Ala Asn Ala Asp Ser Gly Leu Leu His Ser Ser Val
 630 635 640

 AAA AGA AGC TGT TCA CAG AAT GAT TCT GAA GAA CCA ACT TTG TCC TTA 2205
 Lys Arg Ser Cys Ser Gln Asn Asp Ser Glu Glu Pro Thr Leu Ser Leu
 645 650 655

 ACT AGC TCT TTT GGG ACA ATT CTG AGG AAA TGT TCT AGA AAT GAA ACA 2253
 Thr Ser Ser Phe Gly Thr Ile Leu Arg Lys Cys Ser Arg Asn Glu Thr
 660 665 670 675

| | | | |
|---|-----|-----|------|
| TGT TCT AAT AAT ACA GTA ATC TCT CAG GAT CTT GAT TAT AAA GAA GCA Cys Ser Asn Asn Thr Val Ile Ser Gln Asp Leu Asp Tyr Lys Glu Ala 680 | 685 | 690 | 2301 |
| AAA TGT AAT AAG GAA AAA CTA CAG TTA TTT ATT ACC CCA GAA GCT GAT Lys Cys Asn Lys Glu Lys Leu Gln Leu Phe Ile Thr Pro Glu Ala Asp 695 | 700 | 705 | 2349 |
| TCT CTG TCA TGC CTG CAG GAA GGA CAG TGT GAA AAT GAT CCA AAA AGC Ser Leu Ser Cys Leu Gln Glu Gly Gln Cys Glu Asn Asp Pro Lys Ser 710 | 715 | 720 | 2397 |
| AAA AAA GTT TCA GAT ATA AAA GAA GAG GTC TTG GCT GCA GCA TGT CAC Lys Lys Val Ser Asp Ile Lys Glu Glu Val Leu Ala Ala Ala Cys His 725 | 730 | 735 | 2445 |
| CCA GTA CAA CAT TCA AAA GTG GAA TAC AGT GAT ACT GAC TTT CAA TCC Pro Val Gln His Ser Lys Val Glu Tyr Ser Asp Thr Asp Phe Gln Ser 740 | 745 | 750 | 2493 |
| CAG AAA AGT CTT TTA TAT GAT CAT GAA AAT GCC AGC ACT CTT ATT TTA Gln Lys Ser Leu Leu Tyr Asp His Glu Asn Ala Ser Thr Leu Ile Leu 760 | 765 | 770 | 2541 |
| ACT CCT ACT TCC AAG GAT GTT CTG TCA AAC CTA GTC ATG ATT TCT AGA Thr Pro Thr Ser Lys Asp Val Leu Ser Asn Leu Val Met Ile Ser Arg 775 | 780 | 785 | 2589 |
| GGC AAA GAA TCA TAC AAA ATG TCA GAC AAG CTC AAA GGT AAC AAT TAT Gly Lys Ser Tyr Lys Met Ser Asp Lys Leu Lys Gly Asn Asn Tyr 790 | 795 | 800 | 2637 |
| GAA TCT GAT GTT GAA TTA ACC AAA AAT ATT CCC ATG GAA AAG AAT CAA Glu Ser Asp Val Glu Leu Thr Lys Asn Ile Pro Met Glu Lys Asn Gln 805 | 810 | 815 | 2685 |
| GAT GTA TGT GCT TTA AAT GAA AAT TAT AAA AAC GTT GAG CTG TTG CCA Asp Val Cys Ala Leu Asn Glu Asn Tyr Lys Asn Val Glu Leu Leu Pro 820 | 825 | 830 | 2733 |
| CCT GAA AAA TAC ATG AGA GTA GCA TCA CCT TCA AGA AAG GTA CAA TTC Pro Glu Lys Tyr Met Arg Val Ala Ser Pro Ser Arg Lys Val Gln Phe 840 | 845 | 850 | 2781 |
| AAC CAA AAC ACA AAT CTA AGA GTA ATC CAA AAA AAT CAA GAA GAA ACT Asn Gln Asn Thr Asn Leu Arg Val Ile Gln Lys Asn Gln Glu Glu Thr 855 | 860 | 865 | 2829 |
| ACT TCA ATT TCA AAA ATA ACT GTC AAT CCA GAC TCT GAA GAA CTT TTC Thr Ser Ile Ser Lys Ile Thr Val Asn Pro Asp Ser Glu Glu Leu Phe 870 | 875 | 880 | 2877 |
| TCA GAC AAT GAG AAT AAT TTT GTC TTC CAA GTA GCT AAT GAA AGG AAT Ser Asp Asn Glu Asn Asn Phe Val Phe Gln Val Ala Asn Glu Arg Asn 885 | 890 | 895 | 2925 |
| AAT CTT GCT TTA GGA AAT ACT AAG GAA CTT CAT GAA ACA GAC TTG ACT Asn Leu Ala Leu Gly Asn Thr Lys Glu Leu His Glu Thr Asp Leu Thr 900 | 905 | 910 | 2973 |
| TGT GTA AAC GAA CCC ATT TTC AAG AAC TCT ACC ATG GTT TTA TAT GGA | | | 3021 |

Cys Val Asn Glu Pro Ile Phe Lys Asn Ser Thr Met Val Leu Tyr Gly
 920 925 930

GAC ACA GGT GAT AAA CAA GCA ACC CAA GTG TCA ATT AAA AAA GAT TTG 3069
 Asp Thr Gly Asp Lys Gln Ala Thr Gln Val Ser Ile Lys Lys Asp Leu
 935 940 945

GTT TAT GTT CTT GCA GAG GAG AAC AAA AAT AGT GTA AAG CAG CAT ATA 3117
 Val Tyr Val Leu Ala Glu Glu Asn Lys Asn Ser Val Lys Gln His Ile
 950 955 960

AAA ATG ACT CTA GGT CAA GAT TTA AAA TCG GAC ATC TCC TTG AAT ATA 3165
 Lys Met Thr Leu Gly Gln Asp Leu Lys Ser Asp Ile Ser Leu Asn Ile
 965 970 975

GAT AAA ATA CCA GAA AAA AAT AAT GAT TAC ATG AAC AAA TGG GCA GGA 3213
 Asp Lys Ile Pro Glu Lys Asn Asn Asp Tyr Met Asn Lys Trp Ala Gly
 980 985 990 995

CTC TTA GGT CCA ATT TCA AAT CAC AGT TTT GGA GGT AGC TTC AGA ACA 3261
 Leu Leu Gly Pro Ile Ser Asn His Ser Phe Gly Gly Ser Phe Arg Thr
 1000 1005 1010

D/ COMX
 GCT TCA AAT AAG GAA ATC AAG CTC TCT GAA CAT AAC ATT AAG AAG AGC 3309
 Ala Ser Asn Lys Glu Ile Lys Leu Ser Glu His Asn Ile Lys Lys Ser
 1015 1020 1025

AAA ATG TTC TTC AAA GAT ATT GAA GAA CAA TAT CCT ACT AGT TTA GCT 3357
 Lys Met Phe Phe Lys Asp Ile Glu Glu Gln Tyr Pro Thr Ser Leu Ala
 1030 1035 1040

TGT GTT GAA ATT GTA AAT ACC TTG GCA TTA GAT AAT CAA AAG AAA CTG 3405
 Cys Val Glu Ile Val Asn Thr Leu Ala Leu Asp Asn Gln Lys Lys Leu
 1045 1050 1055

AGC AAG CCT CAG TCA ATT AAT ACT GTA TCT GCA CAT TTA CAG AGT AGT 3453
 Ser Lys Pro Gln Ser Ile Asn Thr Val Ser Ala His Leu Gln Ser Ser
 1060 1065 1070 1075

GTA GTT GTT TCT GAT TGT AAA AAT AGT CAT ATA ACC CCT CAG ATG TTA 3501
 Val Val Val Ser Asp Cys Lys Asn Ser His Ile Thr Pro Gln Met Leu
 1080 1085 1090

TTT TCC AAG CAG GAT TTT AAT TCA AAC CAT AAT TTA ACA CCT AGC CAA 3549
 Phe Ser Lys Gln Asp Phe Asn Ser Asn His Asn Leu Thr Pro Ser Gln
 1095 1100 1105

AAG GCA GAA ATT ACA GAA CTT TCT ACT ATA TTA GAA GAA TCA GGA AGT 3597
 Lys Ala Glu Ile Thr Glu Leu Ser Thr Ile Leu Glu Ser Gly Ser
 1110 1115 1120

CAG TTT GAA TTT ACT CAG TTT AGA AAR CCA AGC TAC ATA TTG CAG AAG 3645
 Gln Phe Glu Phe Thr Gln Phe Arg Xaa Pro Ser Tyr Ile Leu Gln Lys
 1125 1130 1135

AGT ACA TTT GAA GTG CCT GAA AAC CAG ATG ACT ATC TTA AAG ACC ACT 3693
 Ser Thr Phe Glu Val Pro Glu Asn Gln Met Thr Ile Leu Lys Thr Thr
 1140 1145 1150 1155

TCT GAG GAA TGC AGA GAT GCT GAT CTT CAT GTC ATA ATG AAT GCC CCA 3741
 Ser Glu Glu Cys Arg Asp Ala Asp Leu His Val Ile Met Asn Ala Pro

| 1160 | 1165 | 1170 | |
|--|------|------|------|
| TCG ATT GGT CAG GTA GAC AGC AGC AAG CAA TTT GAA GGT ACA GTT GAA Ser Ile Gly Gln Val Asp Ser Ser Lys Gln Phe Glu Gly Thr Val Glu 1175 | 1180 | 1185 | 3789 |
| ATT AAA CGG AAG TTT GCT GGC CTG TTG AAA AAT GAC TGT AAC AAA AGT Ile Lys Arg Lys Phe Ala Gly Leu Leu Lys Asn Asp Cys Asn Lys Ser 1190 | 1195 | 1200 | 3837 |
| GCT TCT GGT TAT TTA ACA GAT GAA AAT GAA GTG GGG TTT AGG GGC TTT Ala Ser Gly Tyr Leu Thr Asp Glu Asn Glu Val Gly Phe Arg Gly Phe 1205 | 1210 | 1215 | 3885 |
| TAT TCT GCT CAT GGC ACA AAA CTG AAT GTT TCT ACT GAA GCT CTG CAA Tyr Ser Ala His Gly Thr Lys Leu Asn Val Ser Thr Glu Ala Leu Gln 1220 | 1225 | 1230 | 3933 |
| AAA GCT GTG AAA CTG TTT AGT GAT ATT GAG AAT ATT AGT GAG GAA ACT Lys Ala Val Lys Leu Phe Ser Asp Ile Glu Asn Ile Ser Glu Glu Thr 1240 | 1245 | 1250 | 3981 |
| TCT GCA GAG GTA CAT CCA ATA AGT TTA TCT TCA AGT AAA TGT CAT GAT Ser Ala Glu Val His Pro Ile Ser Leu Ser Ser Lys Cys His Asp 1255 | 1260 | 1265 | 4029 |
| TCT GTC GTT TCA ATG TTT AAG ATA GAA AAT CAT AAT GAT AAA ACT GTA Ser Val Val Ser Met Phe Lys Ile Glu Asn His Asn Asp Lys Thr Val 1270 | 1275 | 1280 | 4077 |
| AGT GAA AAA AAT AAT AAA TGC CAA CTG ATA TTA CAA AAT AAT ATT GAA Ser Glu Lys Asn Asn Lys Cys Gln Leu Ile Leu Gln Asn Asn Ile Glu 1285 | 1290 | 1295 | 4125 |
| ATG ACT ACT GGC ACT TTT GTT GAA GAA ATT ACT GAA AAT TAC AAG AGA Met Thr Thr Gly Thr Phe Val Glu Glu Ile Thr Glu Asn Tyr Lys Arg 1300 | 1305 | 1310 | 4173 |
| AAT ACT GAA AAT GAA GAT AAC AAA TAT ACT GCT GCC AGT AGA AAT TCT Asn Thr Glu Asn Glu Asp Asn Lys Tyr Thr Ala Ala Ser Arg Asn Ser 1320 | 1325 | 1330 | 4221 |
| CAT AAC TTA GAA TTT GAT GGC AGT GAT TCA AGT AAA AAT GAT ACT GTT His Asn Leu Glu Phe Asp Gly Ser Asp Ser Ser Lys Asn Asp Thr Val 1335 | 1340 | 1345 | 4269 |
| TGT ATT CAT AAA GAT GAA ACG GAC TTG CTA TTT ACT GAT CAG CAC AAC Cys Ile His Lys Asp Glu Thr Asp Leu Leu Phe Thr Asp Gln His Asn 1350 | 1355 | 1360 | 4317 |
| ATA TGT CTT AAA TTA TCT GGC CAG TTT ATG AAG GAG GGA AAC ACT CAG Ile Cys Leu Lys Leu Ser Gly Gln Phe Met Lys Glu Gly Asn Thr Gln 1365 | 1370 | 1375 | 4365 |
| ATT AAA GAA GAT TTG TCA GAT TTA ACT TTT TTG GAA GTT GCG AAA GCT Ile Lys Glu Asp Leu Ser Asp Leu Thr Phe Leu Glu Val Ala Lys Ala 1380 | 1385 | 1390 | 4413 |
| CAA GAA GCA TGT CAT GGT AAT ACT TCA AAT AAA GAA CAG TTA ACT GCT Gln Glu Ala Cys His Gly Asn Thr Ser Asn Lys Glu Gln Leu Thr Ala 1400 | 1405 | 1410 | 4461 |

ACT AAA ACG GAG CAA AAT ATA AAA GAT TTT GAG ACT TCT GAT ACA TTT 4509
 Thr Lys Thr Glu Gln Asn Ile Lys Asp Phe Glu Thr Ser Asp Thr Phe
 1415 1420 1425

TTT CAG ACT GCA AGT GGG AAA AAT ATT AGT GTC GCC AAA GAG TCA TTT 4557
 Phe Gln Thr Ala Ser Gly Lys Asn Ile Ser Val Ala Lys Glu Ser Phe
 1430 1435 1440

AAT AAA ATT GTA AAT TTC TTT GAT CAG AAA CCA GAA GAA TTG CAT AAC 4605
 Asn Lys Ile Val Asn Phe Phe Asp Gln Lys Pro Glu Glu Leu His Asn
 1445 1450 1455

TTT TCC TTA AAT TCT GAA TTA CAT TCT GAC ATA AGA AAG AAC AAA ATG 4653
 Phe Ser Leu Asn Ser Glu Leu His Ser Asp Ile Arg Lys Asn Lys Met
 1460 1465 1470 1475

GAC ATT CTA AGT TAT GAG GAA ACA GAC ATA GTT AAA CAC AAA ATA CTG 4701
 Asp Ile Leu Ser Tyr Glu Glu Thr Asp Ile Val Lys His Lys Ile Leu
 1480 1485 1490

AAA GAA AGT GTC CCA GTT GGT ACT GGA AAT CAA CTA GTG ACC TTC CAG 4749
 Lys Glu Ser Val Pro Val Gly Thr Gly Asn Gln Leu Val Thr Phe Gln
 1495 1500 1505

GGA CAA CCC GAA CGT GAT GAA AAG ATC AAA GAA CCT ACT CTG TTG GGT 4797
 Gly Gln Pro Glu Arg Asp Glu Lys Ile Lys Glu Pro Thr Leu Leu Gly
 1510 1515 1520

TTT CAT ACA GCT AGC GGG AAA AAA GTT AAA ATT GCA AAG GAA TCT TTG 4845
 Phe His Thr Ala Ser Gly Lys Val Lys Ile Ala Lys Glu Ser Leu
 1525 1530 1535

GAC AAA GTG AAA AAC CTT TTT GAT GAA AAA GAG CAA GGT ACT AGT GAA 4893
 Asp Lys Val Lys Asn Leu Phe Asp Glu Lys Glu Gln Gly Thr Ser Glu
 1540 1545 1550 1555

ATC ACC AGT TTT AGC CAT CAA TGG GCA AAG ACC CTA AAG TAC AGA GAG 4941
 Ile Thr Ser Phe Ser His Gln Trp Ala Lys Thr Leu Lys Tyr Arg Glu
 1560 1565 1570

GCC TGT AAA GAC CTT GAA TTA GCA TGT GAG ACC ATT GAG ATC ACA GCT 4989
 Ala Cys Lys Asp Leu Glu Leu Ala Cys Glu Thr Ile Glu Ile Thr Ala
 1575 1580 1585

GCC CCA AAG TGT AAA GAA ATG CAG AAT TCT CTC AAT AAT GAT AAA AAC 5037
 Ala Pro Lys Cys Lys Glu Met Gln Asn Ser Leu Asn Asn Asp Lys Asn
 1590 1595 1600

CTT GTT TCT ATT GAG ACT GTG GTG CCA CCT AAG CTC TTA AGT GAT AAT 5085
 Leu Val Ser Ile Glu Thr Val Val Pro Pro Lys Leu Leu Ser Asp Asn
 1605 1610 1615

TTA TGT AGA CAA ACT GAA AAT CTC AAA ACA TCA AAA AGT ATC TTT TTG 5133
 Leu Cys Arg Gln Thr Glu Asn Leu Lys Thr Ser Lys Ser Ile Phe Leu
 1620 1625 1630 1635

AAA GTT AAA GTA CAT GAA AAT GTA GAA AAA GAA ACA GCA AAA AGT CCT 5181
 Lys Val Lys Val His Glu Asn Val Glu Lys Glu Thr Ala Lys Ser Pro
 1640 1645 1650

| | |
|---|------|
| GCA ACT TGT TAC ACA AAT CAG TCC CCT TAT TCA GTC ATT GAA AAT TCA Ala Thr Cys Tyr Thr Asn Gln Ser Pro Tyr Ser Val Ile Glu Asn Ser 1655 1660 1665 | 5229 |
| GCC TTA GCT TTT TAC ACA AGT TGT AGT AGA AAA ACT TCT GTG AGT CAG Ala Leu Ala Phe Tyr Thr Ser Cys Ser Arg Lys Thr Ser Val Ser Gln 1670 1675 1680 | 5277 |
| ACT TCA TTA CTT GAA GCA AAA AAA TGG CTT AGA GAA GGA ATA TTT GAT Thr Ser Leu Leu Glu Ala Lys Lys Trp Leu Arg Glu Gly Ile Phe Asp 1685 1690 1695 | 5325 |
| GGT CAA CCA GAA AGA ATA AAT ACT GCA GAT TAT GTA GGA AAT TAT TTG Gly Gln Pro Glu Arg Ile Asn Thr Ala Asp Tyr Val Gly Asn Tyr Leu 1700 1705 1710 1715 | 5373 |
| TAT GAA AAT AAT TCA AAC AGT ACT ATA GCT GAA AAT GAC AAA AAT CAT Tyr Glu Asn Asn Ser Asn Ser Thr Ile Ala Glu Asn Asp Lys Asn His 1720 1725 1730 | 5421 |
| CTC TCC GAA AAA CAA GAT ACT TAT TTA AGT AAC AGT AGC ATG TCT AAC Leu Ser Glu Lys Gln Asp Thr Tyr Leu Ser Asn Ser Met Ser Asn 1735 1740 1745 | 5469 |
| AGC TAT TCC TAC CAT TCT GAT GAG GTA TAT AAT GAT TCA GGA TAT CTC Ser Tyr Ser Tyr His Ser Asp Glu Val Tyr Asn Asp Ser Gly Tyr Leu 1750 1755 1760 | 5517 |
| TCA AAA AAT AAA CTT GAT TCT GGT ATT GAG CCA GTA TTG AAG AAT GTT Ser Lys Asn Lys Leu Asp Ser Gly Ile Glu Pro Val Leu Lys Asn Val 1765 1770 1775 | 5565 |
| GAA GAT CAA AAA AAC ACT AGT TTT TCC AAA GTA ATA TCC AAT GTA AAA Glu Asp Gln Lys Asn Thr Ser Phe Ser Lys Val Ile Ser Asn Val Lys 1780 1785 1790 1795 | 5613 |
| GAT GCA AAT GCA TAC CCA CAA ACT GTA AAT GAA GAT ATT TGC GTT GAG Asp Ala Asn Ala Tyr Pro Gln Thr Val Asn Glu Asp Ile Cys Val Glu 1800 1805 1810 | 5661 |
| GAA CTT GTG ACT AGC TCT TCA CCC TGC AAA AAT AAA AAT GCA GCC ATT Glu Leu Val Thr Ser Ser Ser Pro Cys Lys Asn Lys Asn Ala Ala Ile 1815 1820 1825 | 5709 |
| AAA TTG TCC ATA TCT AAT AGT AAT AAT TTT GAG GTA GGG CCA CCT GCA Lys Leu Ser Ile Ser Asn Ser Asn Asn Phe Glu Val Gly Pro Pro Ala 1830 1835 1840 | 5757 |
| TTT AGG ATA GCC AGT GGT AAA ATC GTT TGT GTT TCA CAT GAA ACA ATT Phe Arg Ile Ala Ser Gly Lys Ile Val Cys Val Ser His Glu Thr Ile 1845 1850 1855 | 5805 |
| AAA AAA GTG AAA GAC ATA TTT ACA GAC AGT TTC AGT AAA GTA ATT AAG Lys Lys Val Lys Asp Ile Phe Thr Asp Ser Phe Ser Lys Val Ile Lys 1860 1865 1870 1875 | 5853 |
| GAA AAC AAC GAG AAT AAA TCA AAA ATT TGC CAA ACG AAA ATT ATG GCA Glu Asn Asn Glu Asn Lys Ser Lys Ile Cys Gln Thr Lys Ile Met Ala 1880 1885 1890 | 5901 |
| GGT TGT TAC GAG GCA TTG GAT GAT TCA GAG GAT ATT CTT CAT AAC TCT | 5949 |

Gly Cys Tyr Glu Ala Leu Asp Asp Ser Glu Asp Ile Leu His Asn Ser
 1895 1900 1905

CTA GAT AAT GAT GAA TGT AGC ACG CAT TCA CAT AAG GTT TTT GCT GAC 5997
 Leu Asp Asn Asp Glu Cys Ser Thr His Ser His Lys Val Phe Ala Asp
 1910 1915 1920

ATT CAG AGT GAA GAA ATT TTA CAA CAT AAC CAA AAT ATG TCT GGA TTG 6045
 Ile Gln Ser Glu Glu Ile Leu Gln His Asn Gln Asn Met Ser Gly Leu
 1925 1930 1935

GAG AAA GTT TCT AAA ATA TCA CCT TGT GAT GTT AGT TTG GAA ACT TCA 6093
 Glu Lys Val Ser Lys Ile Ser Pro Cys Asp Val Ser Leu Glu Thr Ser
 1940 1945 1950 1955

GAT ATA TGT AAA TGT AGT ATA GGG AAG CTT CAT AAG TCA GTC TCA TCT 6141
 Asp Ile Cys Lys Cys Ser Ile Gly Lys Leu His Lys Ser Val Ser Ser
 1960 1965 1970

GCA AAT ACT TGT GGG ATT TTT AGC ACA GCA AGT GGA AAA TCT GTC CAG 6189
 Ala Asn Thr Cys Gly Ile Phe Ser Thr Ala Ser Gly Lys Ser Val Gln
 1975 1980 1985

*D
CK*
 GTA TCA GAT GCT TCA TTA CAA AAC GCA AGA CAA GTG TTT TCT GAA ATA 6237
 Val Ser Asp Ala Ser Leu Gln Asn Ala Arg Gln Val Phe Ser Glu Ile
 1990 1995 2000

GAA GAT AGT ACC AAG CAA GTC TTT TCC AAA GTA TTG TTT AAA AGT AAC 6285
 Glu Asp Ser Thr Lys Gln Val Phe Ser Lys Val Leu Phe Lys Ser Asn
 2005 2010 2015

GAA CAT TCA GAC CAG CTC ACA AGA GAA AAT ACT GCT ATA CGT ACT 6333
 Glu His Ser Asp Gln Leu Thr Arg Glu Asn Thr Ala Ile Arg Thr
 2020 2025 2030 2035

CCA GAA CAT TTA ATA TCC CAA AAA GGC TTT TCA TAT AAT GTG GTA AAT 6381
 Pro Glu His Leu Ile Ser Gln Lys Gly Phe Ser Tyr Asn Val Val Asn
 2040 2045 2050

TCA TCT GCT TTC TCT GGA TTT AGT ACA GCA AGT GGA AAG CAA GTT TCC 6429
 Ser Ser Ala Phe Ser Gly Phe Ser Thr Ala Ser Gly Lys Gln Val Ser
 2055 2060 2065

ATT TTA GAA AGT TCC TTA CAC AAA GTT AAG GGA GTG TTA GAG GAA TTT 6477
 Ile Leu Glu Ser Ser Leu His Lys Val Lys Gly Val Leu Glu Glu Phe
 2070 2075 2080

GAT TTA ATC AGA ACT GAG CAT AGT CTT CAC TAT TCA CCT ACG TCT AGA 6525
 Asp Leu Ile Arg Thr Glu His Ser Leu His Tyr Ser Pro Thr Ser Arg
 2085 2090 2095

CAA AAT GTA TCA AAA ATA CTT CCT CGT GTT GAT AAG AGA AAC CCA GAG 6573
 Gln Asn Val Ser Lys Ile Leu Pro Arg Val Asp Lys Arg Asn Pro Glu
 2100 2105 2110 2115

CAC TGT GTA AAC TCA GAA ATG GAA AAA ACC TGC AGT AAA GAA TTT AAA 6621
 His Cys Val Asn Ser Glu Met Glu Lys Thr Cys Ser Lys Glu Phe Lys
 2120 2125 2130

TTA TCA AAT AAC TTA AAT GTT GAA GGT GGT TCT TCA GAA AAT AAT CAC 6669
 Leu Ser Asn Asn Leu Asn Val Glu Gly Ser Ser Glu Asn Asn His

2135 2140 2145

| | | |
|--|--|------|
| TCT ATT AAA GTT TCT CCA TAT CTC TCT CAA TTT CAA CAA GAC AAA CAA Ser Ile Lys Val Ser Pro Tyr Leu Ser Gln Phe Gln Gln Asp Lys Gln | | 6717 |
| 2150 2155 2160 | | |
| CAG TTG GTA TTA GGA ACC AAA GTC TCA CTT GTT GAG AAC ATT CAT GTT Gln Leu Val Leu Gly Thr Lys Val Ser Leu Val Glu Asn Ile His Val | | 6765 |
| 2165 2170 2175 | | |
| TTG GGA AAA GAA CAG GCT TCA CCT AAA AAC GTA AAA ATG GAA ATT GGT Leu Gly Lys Glu Gln Ala Ser Pro Lys Asn Val Lys Met Glu Ile Gly | | 6813 |
| 2180 2185 2190 2195 | | |
| AAA ACT GAA ACT TTT TCT GAT GTT CCT GTG AAA ACA AAT ATA GAA GTT Lys Thr Glu Thr Phe Ser Asp Val Pro Val Lys Thr Asn Ile Glu Val | | 6861 |
| 2200 2205 2210 | | |
| TGT TCT ACT TAC TCC AAA GAT TCA GAA AAC TAC TTT GAA ACA GAA GCA Cys Ser Thr Tyr Ser Lys Asp Ser Glu Asn Tyr Phe Glu Thr Glu Ala | | 6909 |
| 2215 2220 2225 | | |
| GTA GAA ATT GCT AAA GCT TTT ATG GAA GAT GAT GAA CTG ACA GAT TCT Val Glu Ile Ala Lys Ala Phe Met Glu Asp Asp Glu Leu Thr Asp Ser | | 6957 |
| 2230 2235 2240 | | |
| AAA CTG CCA AGT CAT GCC ACA CAT TCT CTT TTT ACA TGT CCC GAA AAT Lys Leu Pro Ser His Ala Thr His Ser Leu Phe Thr Cys Pro Glu Asn | | 7005 |
| 2245 2250 2255 | | |
| GAG GAA ATG GTT TTG TCA AAT TCA AGA ATT GGA AAA AGA AGA GGA GAG Glu Glu Met Val Leu Ser Asn Ser Arg Ile Gly Lys Arg Arg Gly Glu | | 7053 |
| 2260 2265 2270 2275 | | |
| CCC CTT ATC TTA GTG GGA GAA CCC TCA ATC AAA AGA AAC TTA TTA AAT Pro Leu Ile Leu Val Gly Glu Pro Ser Ile Lys Arg Asn Leu Leu Asn | | 7101 |
| 2280 2285 2290 | | |
| GAA TTT GAC AGG ATA ATA GAA AAT CAA GAA AAA TCC TTA AAG GCT TCA Glu Phe Asp Arg Ile Ile Glu Asn Gln Glu Lys Ser Leu Lys Ala Ser | | 7149 |
| 2295 2300 2305 | | |
| AAA AGC ACT CCA GAT GGC ACA ATA AAA GAT CGA AGA TTG TTT ATG CAT Lys Ser Thr Pro Asp Gly Thr Ile Lys Asp Arg Arg Leu Phe Met His | | 7197 |
| 2310 2315 2320 | | |
| CAT GTT TCT TTA GAG CCG ATT ACC TGT GTA CCC TTT CGC ACA ACT AAG His Val Ser Leu Glu Pro Ile Thr Cys Val Pro Phe Arg Thr Thr Lys | | 7245 |
| 2325 2330 2335 | | |
| GAA CGT CAA GAG ATA CAG AAT CCA AAT TTT ACC GCA CCT GGT CAA GAA Glu Arg Gln Glu Ile Gln Asn Pro Asn Phe Thr Ala Pro Gly Gln Glu | | 7293 |
| 2340 2345 2350 2355 | | |
| TTT CTG TCT AAA TCT CAT TTG TAT GAA CAT CTG ACT TTG GAA AAA TCT Phe Leu Ser Lys Ser His Leu Tyr Glu His Leu Thr Leu Glu Lys Ser | | 7341 |
| 2360 2365 2370 | | |
| TCA AGC AAT TTA GCA GTT TCA GGA CAT CCA TTT TAT CAA GTT TCT GCT Ser Ser Asn Leu Ala Val Ser Gly His Pro Phe Tyr Gln Val Ser Ala | | 7389 |
| 2375 2380 2385 | | |

ACA AGA AAT GAA AAA ATG AGA CAC TTG ATT ACT ACA GGC AGA CCA ACC 7437
Thr Arg Asn Glu Lys Met Arg His Leu Ile Thr Thr Gly Arg Pro Thr
2390 2395 2400

AAA GTC TTT GTT CCA CCT TTT AAA ACT AAA TCA CAT TTT CAC AGA GTT 7485
Lys Val Phe Val Pro Pro Phe Lys Thr Lys Ser His Phe His Arg Val
2405 2410 2415

GAA CAG TGT GTT AGG AAT ATT AAC TTG GAG GAA AAC AGA CAA AAG CAA 7533
Glu Gln Cys Val Arg Asn Ile Asn Leu Glu Glu Asn Arg Gln Lys Gln
2420 2425 2430 2435

AAC ATT GAT GGA CAT GGC TCT GAT GAT AGT AAA AAT AAG ATT AAT GAC 7581
Asn Ile Asp Gly His Gly Ser Asp Asp Ser Lys Asn Lys Ile Asn Asp
2440 2445 2450

AAT GAG ATT CAT CAG TTT AAC AAA AAC AAC TCC AAT CAA GCA GCA GCT 7629
Asn Glu Ile His Gln Phe Asn Lys Asn Asn Ser Asn Gln Ala Ala Ala
2455 2460 2465


GTA ACT TTC ACA AAG TGT GAA GAA GAA CCT TTA GAT TTA ATT ACA AGT 7677
Val Thr Phe Thr Lys Cys Glu Glu Pro Leu Asp Leu Ile Thr Ser
2470 2475 2480

CTT CAG AAT GCC AGA GAT ATA CAG GAT ATG CGA ATT AAG AAG AAA CAA 7725
Leu Gln Asn Ala Arg Asp Ile Gln Asp Met Arg Ile Lys Lys Lys Gln
2485 2490 2495

AGG CAA CGC GTC TTT CCA CAG CCA GGC AGT CTG TAT CTT GCA AAA ACA 7773
Arg Gln Arg Val Phe Pro Gln Pro Gly Ser Leu Tyr Leu Ala Lys Thr
2500 2505 2510 2515

TCC ACT CTG CCT CGA ATC TCT CTG AAA GCA GCA GTA GGA GGC CAA GTT 7821
Ser Thr Leu Pro Arg Ile Ser Leu Lys Ala Ala Val Gly Gly Gln Val
2520 2525 2530

CCC TCT GCG TGT TCT CAT AAA CAG CTG TAT ACG TAT GGC GTT TCT AAA 7869
Pro Ser Ala Cys Ser His Lys Gln Leu Tyr Thr Tyr Gly Val Ser Lys
2535 2540 2545

CAT TGC ATA AAA ATT AAC AGC AAA AAT GCA GAG TCT TTT CAG TTT CAC 7917
His Cys Ile Lys Ile Asn Ser Lys Asn Ala Glu Ser Phe Gln Phe His
2550 2555 2560

ACT GAA GAT TAT TTT GGT AAG GAA AGT TTA TGG ACT GGA AAA GGA ATA 7965
Thr Glu Asp Tyr Phe Gly Lys Glu Ser Leu Trp Thr Gly Lys Gly Ile
2565 2570 2575

CAG TTG GCT GAT GGT GGA TGG CTC ATA CCC TCC AAT GAT GGA AAG GCT 8013
Gln Leu Ala Asp Gly Gly Trp Leu Ile Pro Ser Asn Asp Gly Lys Ala
2580 2585 2590 2595

GGA AAA GAA GAA TTT TAT AGG GCT CTG TGT GAC ACT CCA GGT GTG GAT 8061
Gly Lys Glu Glu Phe Tyr Arg Ala Leu Cys Asp Thr Pro Gly Val Asp
2600 2605 2610

CCA AAG CTT ATT TCT AGA ATT TGG GTT TAT AAT CAC TAT AGA TGG ATC 8109
Pro Lys Leu Ile Ser Arg Ile Trp Val Tyr Asn His Tyr Arg Trp Ile
2615 2620 2625

| | | | |
|--|------|------|------|
| ATA TGG AAA CTG GCA GCT ATG GAA TGT GCC TTT CCT AAG GAA TTT GCT Ile Trp Lys Leu Ala Ala Met Glu Cys Ala Phe Pro Lys Glu Phe Ala 2630 | 2635 | 2640 | 8157 |
| AAT AGA TGC CTA AGC CCA GAA AGG GTG CTT CTT CAA CTA AAA TAC AGA Asn Arg Cys Leu Ser Pro Glu Arg Val Leu Leu Gln Leu Lys Tyr Arg 2645 | 2650 | 2655 | 8205 |
| TAT GAT ACG GAA ATT GAT AGA AGC AGA AGA TCG GCT ATA AAA AAG ATA Tyr Asp Thr Glu Ile Asp Arg Ser Arg Arg Ser Ala Ile Lys Lys Ile 2660 | 2665 | 2670 | 2675 |
| ATG GAA AGG GAT GAC ACA GCT GCA AAA ACA CTT GTT CTC TGT GTT TCT Met Glu Arg Asp Asp Thr Ala Ala Lys Thr Leu Val Leu Cys Val Ser 2680 | 2685 | 2690 | 8301 |
| GAC ATA ATT TCA TTG AGC GCA AAT ATA TCT GAA ACT TCT AGC AAT AAA Asp Ile Ile Ser Leu Ser Ala Asn Ile Ser Glu Thr Ser Ser Asn Lys 2695 | 2700 | 2705 | 8349 |
| ACT AGT AGT GCA GAT ACC CAA AAA GTG GCC ATT ATT GAA CTT ACA GAT Thr Ser Ser Ala Asp Thr Gln Lys Val Ala Ile Ile Glu Leu Thr Asp 2710 | 2715 | 2720 | 8397 |
| GGG TGG TAT GCT GTT AAG GCC CAG TTA GAT CCT CCC CTC TTA GCT GTC Gly Trp Tyr Ala Val Lys Ala Gln Leu Asp Pro Pro Leu Leu Ala Val 2725 | 2730 | 2735 | 8445 |
| TTA AAG AAT GGC AGA CTG ACA GTT GGT CAG AAG ATT ATT CTT CAT GGA Leu Lys Asn Gly Arg Leu Thr Val Gly Gln Lys Ile Ile Leu His Gly 2740 | 2745 | 2750 | 2755 |
| GCA GAA CTG GTG GGC TCT CCT GAT GCC TGT ACA CCT CTT GAA GCC CCA Ala Glu Leu Val Gly Ser Pro Asp Ala Cys Thr Pro Leu Glu Ala Pro 2760 | 2765 | 2770 | 8541 |
| GAA TCT CTT ATG TTA AAG ATT TCT GCT AAC AGT ACT CGG CCT GCT CGC Glu Ser Leu Met Leu Lys Ile Ser Ala Asn Ser Thr Arg Pro Ala Arg 2775 | 2780 | 2785 | 8589 |
| TGG TAT ACC AAA CTT GGA TTC TTT CCT GAC CCT AGA CCT TTT CCT CTG Trp Tyr Thr Lys Leu Gly Phe Phe Pro Asp Pro Arg Pro Phe Pro Leu 2790 | 2795 | 2800 | 8637 |
| CCC TTA TCA TCG CTT TTC AGT GAT GGA GGA AAT GTT GGT TGT GTT GAT Pro Leu Ser Ser Leu Phe Ser Asp Gly Gly Asn Val Gly Cys Val Asp 2805 | 2810 | 2815 | 8685 |
| GTA ATT ATT CAA AGA GCA TAC CCT ATA CAG TGG ATG GAG AAG ACA TCA Val Ile Ile Gln Arg Ala Tyr Pro Ile Gln Trp Met Glu Lys Thr Ser 2820 | 2825 | 2830 | 2835 |
| TCT GGA TTA TAC ATA TTT CGC AAT GAA AGA GAG GAA GAA AAG GAA GCA Ser Gly Leu Tyr Ile Phe Arg Asn Glu Arg Glu Glu Glu Lys Glu Ala 2840 | 2845 | 2850 | 8781 |
| GCA AAA TAT GTG GAG GCC CAA AAG AGA CTA GAA GCC TTA TTC ACT Ala Lys Tyr Val Glu Ala Gln Gln Lys Arg Leu Glu Ala Leu Phe Thr 2855 | 2860 | 2865 | 8829 |
| AAA ATT CAG GAG GAA TTT GAA GAA CAT GAA GAA AAC ACA ACA AAA CCA | | | 8877 |

Lys Ile Gln Glu Glu Phe Glu Glu His Glu Glu Asn Thr Thr Lys Pro
 2870 2875 2880

TAT TTA CCA TCA CGT GCA CTA ACA AGA CAG CAA GTT CGT GCT TTG CAA 8925
 Tyr Leu Pro Ser Arg Ala Leu Thr Arg Gln Gln Val Arg Ala Leu Gln
 2885 2890 2895

GAT GGT GCA GAG CTT TAT GAA GCA GTG AAG AAT GCA GCA GAC CCA GCT 8973
 Asp Gly Ala Glu Leu Tyr Glu Ala Val Lys Asn Ala Ala Asp Pro Ala
 2900 2905 2910 2915

TAC CTT GAG GGT TAT TTC AGT GAA GAG CAG TTA AGA GCC TTG AAT AAT 9021
 Tyr Leu Glu Gly Tyr Phe Ser Glu Glu Gln Leu Arg Ala Leu Asn Asn
 2920 2925 2930

CAC AGG CAA ATG TTG AAT GAT AAG AAA CAA GCT CAG ATC CAG TTG GAA 9069
 His Arg Gln Met Leu Asn Asp Lys Lys Gln Ala Gln Ile Gln Leu Glu
 2935 2940 2945

ATT AGG AAG GCC ATG GAA TCT GCT GAA CAA AAG GAA CAA GGT TTA TCA 9117
 Ile Arg Lys Ala Met Glu Ser Ala Glu Gln Lys Glu Gln Gly Leu Ser
 2950 2955 2960

AGG GAT GTC ACA ACC GTG TGG AAG TTG CGT ATT GTA AGC TAT TCA AAA 9165
 Arg Asp Val Thr Thr Val Trp Lys Leu Arg Ile Val Ser Tyr Ser Lys
 2965 2970 2975

AAA GAA AAA GAT TCA GTT ATA CTG AGT ATT TGG CGT CCA TCA TCA GAT 9213
 Lys Glu Lys Asp Ser Val Ile Leu Ser Ile Trp Arg Pro Ser Ser Asp
 2980 2985 2990 2995

TTA TAT TCT CTG TTA ACA GAA GGA AAG AGA TAC AGA ATT TAT CAT CTT 9261
 Leu Tyr Ser Leu Leu Thr Glu Gly Lys Arg Tyr Arg Ile Tyr His Leu
 3000 3005 3010

GCA ACT TCA AAA TCT AAA AGT AAA TCT GAA AGA GCT AAC ATA CAG TTA 9309
 Ala Thr Ser Lys Ser Lys Ser Glu Arg Ala Asn Ile Gln Leu
 3015 3020 3025

GCA GCG ACA AAA AAA ACT CAG TAT CAA CAA CTA CCG GTT TCA GAT GAA 9357
 Ala Ala Thr Lys Lys Thr Gln Tyr Gln Gln Leu Pro Val Ser Asp Glu
 3030 3035 3040

ATT TTA TTT CAG ATT TAC CAG CCA CGG GAG CCC CTT CAC TTC AGC AAA 9405
 Ile Leu Phe Gln Ile Tyr Gln Pro Arg Glu Pro Leu His Phe Ser Lys
 3045 3050 3055

TTT TTA GAT CCA GAC TTT CAG CCA TCT TGT TCT GAG GTG GAC CTA ATA 9453
 Phe Leu Asp Pro Asp Phe Gln Pro Ser Cys Ser Glu Val Asp Leu Ile
 3060 3065 3070 3075

GGA TTT GTC GTT TCT GTG AAA AAA ACA GGA CTT GCC CCT TTC GTC 9501
 Gly Phe Val Val Ser Val Val Lys Lys Thr Gly Leu Ala Pro Phe Val
 3080 3085 3090

TAT TTG TCA GAC GAA TGT TAC AAT TTA CTG GCA ATA AAG TTT TGG ATA 9549
 Tyr Leu Ser Asp Glu Cys Tyr Asn Leu Leu Ala Ile Lys Phe Trp Ile
 3095 3100 3105

GAC CTT AAT GAG GAC ATT ATT AAG CCT CAT ATG TTA ATT GCT GCA AGC 9597
 Asp Leu Asn Glu Asp Ile Ile Lys Pro His Met Leu Ile Ala Ala Ser

| 3110 | 3115 | 3120 | |
|--|------|------|-------|
| AAC CTC CAG TGG CGA CCA GAA TCC AAA TCA GGC CTT CTT ACT TTA TTT Asn Leu Gln Trp Arg Pro Glu Ser Lys Ser Gly Leu Leu Thr Leu Phe 3125 | 3130 | 3135 | 9645 |
| GCT GGA GAT TTT TCT GTG TTT TCT GCT AGT CCA AAA GAG GGC CAC TTT Ala Gly Asp Phe Ser Val Phe Ser Ala Ser Pro Lys Glu Gly His Phe 3140 | 3145 | 3150 | 9693 |
| CAA GAG ACA TTC AAC AAA ATG AAA AAT ACT GTT GAG AAT ATT GAC ATA Gln Glu Thr Phe Asn Lys Met Lys Asn Thr Val Glu Asn Ile Asp Ile 3160 | 3165 | 3170 | 9741 |
| CTT TGC AAT GAA GCA GAA AAC AAG CTT ATG CAT ATA CTG CAT GCA AAT Leu Cys Asn Glu Ala Glu Asn Lys Leu Met His Ile Leu His Ala Asn 3175 | 3180 | 3185 | 9789 |
| GAT CCC AAG TGG TCC ACC CCA ACT AAA GAC TGT ACT TCA GGG CCG TAC Asp Pro Lys Trp Ser Thr Pro Thr Lys Asp Cys Thr Ser Gly Pro Tyr 3190 | 3195 | 3200 | 9837 |
| ACT GCT CAA ATC ATT CCT GGT ACA GGA AAC AAG CTT CTG ATG TCT TCT Thr Ala Gln Ile Ile Pro Gly Thr Gly Asn Lys Leu Leu Met Ser Ser 3205 | 3210 | 3215 | 9885 |
| CCT AAT TGT GAG ATA TAT TAT CAA AGT CCT TTA TCA CTT TGT ATG GCC Pro Asn Cys Glu Ile Tyr Tyr Gln Ser Pro Leu Ser Leu Cys Met Ala 3220 | 3225 | 3230 | 9933 |
| AAA AGG AAG TCT GTT TCC ACA CCT GTC TCA GCC CAG ATG ACT TCA AAG Lys Arg Lys Ser Val Ser Thr Pro Val Ser Ala Gln Met Thr Ser Lys 3240 | 3245 | 3250 | 9981 |
| TCT TGT AAA GGG GAG AAA GAG ATT GAT GAC CAA AAG AAC TGC AAA AAG Ser Cys Lys Gly Glu Lys Glu Ile Asp Asp Gln Lys Asn Cys Lys Lys 3255 | 3260 | 3265 | 10029 |
| AGA AGA GCC TTG GAT TTC TTG AGT AGA CTG CCT TTA CCT CCA CCT GTT Arg Arg Ala Leu Asp Phe Leu Ser Arg Leu Pro Leu Pro Pro Val 3270 | 3275 | 3280 | 10077 |
| AGT CCC ATT TGT ACA TTT GTT TCT CCG GCT GCA CAG AAG GCA TTT CAG Ser Pro Ile Cys Thr Phe Val Ser Pro Ala Ala Gln Lys Ala Phe Gln 3285 | 3290 | 3295 | 10125 |
| CCA CCA AGG AGT TGT GGC ACC AAA TAC GAA ACA CCC ATA AAG AAA AAA Pro Pro Arg Ser Cys Gly Thr Lys Tyr Glu Thr Pro Ile Lys Lys Lys 3300 | 3305 | 3310 | 10173 |
| GAA CTG AAT TCT CCT CAG ATG ACT CCA TTT AAA AAA TTC AAT GAA ATT Glu Leu Asn Ser Pro Gln Met Thr Pro Phe Lys Lys Phe Asn Glu Ile 3320 | 3325 | 3330 | 10221 |
| TCT CTT TTG GAA AGT AAT TCA ATA GCT GAC GAA GAA CTT GCA TTG ATA Ser Leu Leu Glu Ser Asn Ser Ile Ala Asp Glu Glu Leu Ala Leu Ile 3335 | 3340 | 3345 | 10269 |
| AAT ACC CAA GCT CTT TTG TCT GGT TCA ACA GGA GAA AAA CAA TTT ATA Asn Thr Gln Ala Leu Leu Ser Gly Ser Thr Gly Glu Lys Gln Phe Ile 3350 | 3355 | 3360 | 10317 |

| | |
|---|-------|
| TCT GTC AGT GAA TCC ACT AGG ACT GCT CCC ACC AGT TCA GAA GAT TAT | 10365 |
| Ser Val Ser Glu Ser Thr Arg Thr Ala Pro Thr Ser Ser Glu Asp Tyr | |
| 3365 3370 3375 | |
| CTC AGA CTG AAA CGA CGT TGT ACT ACA TCT CTG ATC AAA GAA CAG GAG | 10413 |
| Leu Arg Leu Lys Arg Arg Cys Thr Thr Ser Leu Ile Lys Glu Gln Glu | |
| 3380 3385 3390 3395 | |
| AGT TCC CAG GCC AGT ACG GAA GAA TGT GAG AAA AAT AAG CAG GAC ACA | 10461 |
| Ser Ser Gln Ala Ser Thr Glu Glu Cys Glu Lys Asn Lys Gln Asp Thr | |
| 3400 3405 3410 | |
| ATT ACA ACT AAA AAA TAT ATC TAA | 10485 |
| Ile Thr Thr Lys Lys Tyr Ile | |
| 3415 | |

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3418 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal

(ix) FEATURE:

(A) NAME / K

- (B) LOCATION: 1132
(D) OTHER INFORMATION: Xaa = Lys (from codon AAR,
R = A or G).

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Pro | Ile | Gly | Ser | Lys | Glu | Arg | Pro | Thr | Phe | Phe | Glu | Ile | Phe | Lys |
| 1 | | | | | 5 | | | | 10 | | | | | 15 | |
| Thr | Arg | Cys | Asn | Lys | Ala | Asp | Leu | Gly | Pro | Ile | Ser | Leu | Asn | Trp | Phe |
| | | | | | 20 | | | 25 | | | | | 30 | | |
| Glu | Glu | Leu | Ser | Ser | Glu | Ala | Pro | Pro | Tyr | Asn | Ser | Glu | Pro | Ala | Glu |
| | | | | | 35 | | | 40 | | | | 45 | | | |
| Glu | Ser | Glu | His | Lys | Asn | Asn | Asn | Tyr | Glu | Pro | Asn | Leu | Phe | Lys | Thr |
| | | | | | 50 | | | 55 | | | 60 | | | | |
| Pro | Gln | Arg | Lys | Pro | Ser | Tyr | Asn | Gln | Leu | Ala | Ser | Thr | Pro | Ile | Ile |
| | | | | | 65 | | | 70 | | | 75 | | | 80 | |
| Phe | Lys | Glu | Gln | Gly | Leu | Thr | Leu | Pro | Leu | Tyr | Gln | Ser | Pro | Val | Lys |
| | | | | | 85 | | | | 90 | | | | 95 | | |
| Glu | Leu | Asp | Lys | Phe | Lys | Leu | Asp | Leu | Gly | Arg | Asn | Val | Pro | Asn | Ser |
| | | | | | 100 | | | 105 | | | | 110 | | | |
| Arg | His | Lys | Ser | Leu | Arg | Thr | Val | Lys | Thr | Lys | Met | Asp | Gln | Ala | Asp |
| | | | | | 115 | | | 120 | | | 125 | | | | |
| Asp | Val | Ser | Cys | Pro | Leu | Leu | Asn | Ser | Cys | Leu | Ser | Glu | Ser | Pro | Val |
| | | | | | 130 | | | 135 | | | 140 | | | | |
| Val | Leu | Gln | Cys | Thr | His | Val | Thr | Pro | Gln | Arg | Asp | Lys | Ser | Val | Val |
| | | | | | 145 | | | 150 | | | 155 | | | 160 | |
| Cys | Gly | Ser | Leu | Phe | His | Thr | Pro | Lys | Phe | Val | Lys | Gly | Arg | Gln | Thr |
| | | | | | 165 | | | | 170 | | | | 175 | | |
| Pro | Lys | His | Ile | Ser | Glu | Ser | Leu | Gly | Ala | Glu | Val | Asp | Pro | Asp | Met |
| | | | | | 180 | | | 185 | | | | 190 | | | |
| Ser | Trp | Ser | Ser | Ser | Leu | Ala | Thr | Pro | Pro | Thr | Leu | Ser | Ser | Thr | Val |

195 200 205
 Leu Ile Val Arg Asn Glu Glu Ala Ser Glu Thr Val Phe Pro His Asp
 210 215 220
 Thr Thr Ala Asn Val Lys Ser Tyr Phe Ser Asn His Asp Glu Ser Leu
 225 230 235 240
 Lys Lys Asn Asp Arg Phe Ile Ala Ser Val Thr Asp Ser Glu Asn Thr
 245 250 255
 Asn Gln Arg Glu Ala Ala Ser His Gly Phe Gly Lys Thr Ser Gly Asn
 260 265 270
 Ser Phe Lys Val Asn Ser Cys Lys Asp His Ile Gly Lys Ser Met Pro
 275 280 285
 Asn Val Leu Glu Asp Glu Val Tyr Glu Thr Val Val Asp Thr Ser Glu
 290 295 300
 Glu Asp Ser Phe Ser Leu Cys Phe Ser Lys Cys Arg Thr Lys Asn Leu
 305 310 315 320
 Gln Lys Val Arg Thr Ser Lys Thr Arg Lys Lys Ile Phe His Glu Ala
 325 330 335
 Asn Ala Asp Glu Cys Glu Lys Ser Lys Asn Gln Val Lys Glu Lys Tyr
 340 345 350
 Ser Phe Val Ser Glu Val Glu Pro Asn Asp Thr Asp Pro Leu Asp Ser
 355 360 365
 Asn Val Ala His Gln Lys Pro Phe Glu Ser Gly Ser Asp Lys Ile Ser
 370 375 380
 Lys Glu Val Val Pro Ser Leu Ala Cys Glu Trp Ser Gln Leu Thr Leu
 385 390 395 400
 Ser Gly Leu Asn Gly Ala Gln Met Glu Lys Ile Pro Leu Leu His Ile
 405 410 415
 Ser Ser Cys Asp Gln Asn Ile Ser Glu Lys Asp Leu Leu Asp Thr Glu
 420 425 430
 Asn Lys Arg Lys Lys Asp Phe Leu Thr Ser Glu Asn Ser Leu Pro Arg
 435 440 445
 Ile Ser Ser Leu Pro Lys Ser Glu Lys Pro Leu Asn Glu Glu Thr Val
 450 455 460
 Val Asn Lys Arg Asp Glu Glu Gln His Leu Glu Ser His Thr Asp Cys
 465 470 475 480
 Ile Leu Ala Val Lys Gln Ala Ile Ser Gly Thr Ser Pro Val Ala Ser
 485 490 495
 Ser Phe Gln Gly Ile Lys Lys Ser Ile Phe Arg Ile Arg Glu Ser Pro
 500 505 510
 Lys Glu Thr Phe Asn Ala Ser Phe Ser Gly His Met Thr Asp Pro Asn
 515 520 525
 Phe Lys Lys Glu Thr Glu Ala Ser Glu Ser Gly Leu Glu Ile His Thr
 530 535 540
 Val Cys Ser Gln Lys Glu Asp Ser Leu Cys Pro Asn Leu Ile Asp Asn
 545 550 555 560
 Gly Ser Trp Pro Ala Thr Thr Gln Asn Ser Val Ala Leu Lys Asn
 565 570 575
 Ala Gly Leu Ile Ser Thr Leu Lys Lys Thr Asn Lys Phe Ile Tyr
 580 585 590
 Ala Ile His Asp Glu Thr Ser Tyr Lys Gly Lys Lys Ile Pro Lys Asp
 595 600 605
 Gln Lys Ser Glu Leu Ile Asn Cys Ser Ala Gln Phe Glu Ala Asn Ala
 610 615 620
 Phe Glu Ala Pro Leu Thr Phe Ala Asn Ala Asp Ser Gly Leu Leu His
 625 630 635 640
 Ser Ser Val Lys Arg Ser Cys Ser Gln Asn Asp Ser Glu Glu Pro Thr
 645 650 655
 Leu Ser Leu Thr Ser Ser Phe Gly Thr Ile Leu Arg Lys Cys Ser Arg
 660 665 670
 Asn Glu Thr Cys Ser Asn Asn Thr Val Ile Ser Gln Asp Leu Asp Tyr
 675 680 685

Lys Glu Ala Lys Cys Asn Lys Glu Lys Leu Gln Leu Phe Ile Thr Pro
 690 695 700
 Glu Ala Asp Ser Leu Ser Cys Leu Gln Glu Gly Gln Cys Glu Asn Asp
 705 710 715 720
 Pro Lys Ser Lys Lys Val Ser Asp Ile Lys Glu Glu Val Leu Ala Ala
 725 730 735
 Ala Cys His Pro Val Gln His Ser Lys Val Glu Tyr Ser Asp Thr Asp
 740 745 750
 Phe Gln Ser Gln Lys Ser Leu Leu Tyr Asp His Glu Asn Ala Ser Thr
 755 760 765
 Leu Ile Leu Thr Pro Thr Ser Lys Asp Val Leu Ser Asn Leu Val Met
 770 775 780
 Ile Ser Arg Gly Lys Glu Ser Tyr Lys Met Ser Asp Lys Leu Lys Gly
 785 790 795 800
 Asn Asn Tyr Glu Ser Asp Val Glu Leu Thr Lys Asn Ile Pro Met Glu
 805 810 815
 Lys Asn Gln Asp Val Cys Ala Leu Asn Glu Asn Tyr Lys Asn Val Glu
 820 825 830
 Leu Leu Pro Pro Glu Lys Tyr Met Arg Val Ala Ser Pro Ser Arg Lys
 835 840 845
 Val Gln Phe Asn Gln Asn Thr Asn Leu Arg Val Ile Gln Lys Asn Gln
 850 855 860
 Glu Glu Thr Thr Ser Ile Ser Lys Ile Thr Val Asn Pro Asp Ser Glu
 865 870 875 880
 Glu Leu Phe Ser Asp Asn Glu Asn Asn Phe Val Phe Gln Val Ala Asn
 885 890 895
 Glu Arg Asn Asn Leu Ala Leu Gly Asn Thr Lys Glu Leu His Glu Thr
 900 905 910
 Asp Leu Thr Cys Val Asn Glu Pro Ile Phe Lys Asn Ser Thr Met Val
 915 920 925
 Leu Tyr Gly Asp Thr Gly Asp Lys Gln Ala Thr Gln Val Ser Ile Lys
 930 935 940
 Lys Asp Leu Val Tyr Val Leu Ala Glu Glu Asn Lys Asn Ser Val Lys
 945 950 955 960
 Gln His Ile Lys Met Thr Leu Gly Gln Asp Leu Lys Ser Asp Ile Ser
 965 970 975
 Leu Asn Ile Asp Lys Ile Pro Glu Lys Asn Asn Asp Tyr Met Asn Lys
 980 985 990
 Trp Ala Gly Leu Leu Gly Pro Ile Ser Asn His Ser Phe Gly Gly Ser
 995 1000 1005
 Phe Arg Thr Ala Ser Asn Lys Glu Ile Lys Leu Ser Glu His Asn Ile
 1010 1015 1020
 Lys Lys Ser Lys Met Phe Phe Lys Asp Ile Glu Glu Gln Tyr Pro Thr
 1025 1030 1035 1040
 Ser Leu Ala Cys Val Glu Ile Val Asn Thr Leu Ala Leu Asp Asn Gln
 1045 1050 1055
 Lys Lys Leu Ser Lys Pro Gln Ser Ile Asn Thr Val Ser Ala His Leu
 1060 1065 1070
 Gln Ser Ser Val Val Ser Asp Cys Lys Asn Ser His Ile Thr Pro
 1075 1080 1085
 Gln Met Leu Phe Ser Lys Gln Asp Phe Asn Ser Asn His Asn Leu Thr
 1090 1095 1100
 Pro Ser Gln Lys Ala Glu Ile Thr Glu Leu Ser Thr Ile Leu Glu Glu
 1105 1110 1115 1120
 Ser Gly Ser Gln Phe Glu Phe Thr Gln Phe Arg Xaa Pro Ser Tyr Ile
 1125 1130 1135
 Leu Gln Lys Ser Thr Phe Glu Val Pro Glu Asn Gln Met Thr Ile Leu
 1140 1145 1150
 Lys Thr Thr Ser Glu Glu Cys Arg Asp Ala Asp Leu His Val Ile Met
 1155 1160 1165
 Asn Ala Pro Ser Ile Gly Gln Val Asp Ser Ser Lys Gln Phe Glu Gly

D
X
CON

1170 1175 1180
 Thr Val Glu Ile Lys Arg Lys Phe Ala Gly Leu Leu Lys Asn Asp Cys
 1185 1190 1195 1200
 Asn Lys Ser Ala Ser Gly Tyr Leu Thr Asp Glu Asn Glu Val Gly Phe
 1205 1210 1215
 Arg Gly Phe Tyr Ser Ala His Gly Thr Lys Leu Asn Val Ser Thr Glu
 1220 1225 1230
 Ala Leu Gln Lys Ala Val Lys Leu Phe Ser Asp Ile Glu Asn Ile Ser
 1235 1240 1245
 Glu Glu Thr Ser Ala Glu Val His Pro Ile Ser Leu Ser Ser Ser Lys
 1250 1255 1260
 Cys His Asp Ser Val Val Ser Met Phe Lys Ile Glu Asn His Asn Asp
 1265 1270 1275 1280
 Lys Thr Val Ser Glu Lys Asn Asn Lys Cys Gln Leu Ile Leu Gln Asn
 1285 1290 1295
 Asn Ile Glu Met Thr Thr Gly Thr Phe Val Glu Glu Ile Thr Glu Asn
 1300 1305 1310
 Tyr Lys Arg Asn Thr Glu Asn Glu Asp Asn Lys Tyr Thr Ala Ala Ser
 1315 1320 1325
 Arg Asn Ser His Asn Leu Glu Phe Asp Gly Ser Asp Ser Ser Lys Asn
 1330 1335 1340
 Asp Thr Val Cys Ile His Lys Asp Glu Thr Asp Leu Leu Phe Thr Asp
 1345 1350 1355 1360
 Gln His Asn Ile Cys Leu Lys Leu Ser Gly Gln Phe Met Lys Glu Gly
 1365 1370 1375
 Asn Thr Gln Ile Lys Glu Asp Leu Ser Asp Leu Thr Phe Leu Glu Val
 1380 1385 1390
 Ala Lys Ala Gln Glu Ala Cys His Gly Asn Thr Ser Asn Lys Glu Gln
 1395 1400 1405
 Leu Thr Ala Thr Lys Thr Glu Gln Asn Ile Lys Asp Phe Glu Thr Ser
 1410 1415 1420
 Asp Thr Phe Phe Gln Thr Ala Ser Gly Lys Asn Ile Ser Val Ala Lys
 1425 1430 1435 1440
 Glu Ser Phe Asn Lys Ile Val Asn Phe Phe Asp Gln Lys Pro Glu Glu
 1445 1450 1455
 Leu His Asn Phe Ser Leu Asn Ser Glu Leu His Ser Asp Ile Arg Lys
 1460 1465 1470
 Asn Lys Met Asp Ile Leu Ser Tyr Glu Glu Thr Asp Ile Val Lys His
 1475 1480 1485
 Lys Ile Leu Lys Glu Ser Val Pro Val Gly Thr Gly Asn Gln Leu Val
 1490 1495 1500
 Thr Phe Gln Gly Gln Pro Glu Arg Asp Glu Lys Ile Lys Glu Pro Thr
 1505 1510 1515 1520
 Leu Leu Gly Phe His Thr Ala Ser Gly Lys Lys Val Lys Ile Ala Lys
 1525 1530 1535
 Glu Ser Leu Asp Lys Val Lys Asn Leu Phe Asp Glu Lys Glu Gln Gly
 1540 1545 1550
 Thr Ser Glu Ile Thr Ser Phe Ser His Gln Trp Ala Lys Thr Leu Lys
 1555 1560 1565
 Tyr Arg Glu Ala Cys Lys Asp Leu Glu Leu Ala Cys Glu Thr Ile Glu
 1570 1575 1580
 Ile Thr Ala Ala Pro Lys Cys Lys Glu Met Gln Asn Ser Leu Asn Asn
 1585 1590 1595 1600
 Asp Lys Asn Leu Val Ser Ile Glu Thr Val Val Pro Pro Lys Leu Leu
 1605 1610 1615
 Ser Asp Asn Leu Cys Arg Gln Thr Glu Asn Leu Lys Thr Ser Lys Ser
 1620 1625 1630
 Ile Phe Leu Lys Val Lys Val His Glu Asn Val Glu Lys Glu Thr Ala
 1635 1640 1645
 Lys Ser Pro Ala Thr Cys Tyr Thr Asn Gln Ser Pro Tyr Ser Val Ile
 1650 1655 1660

Glu Asn Ser Ala Leu Ala Phe Tyr Thr Ser Cys Ser Arg Lys Thr Ser
 1665 1670 1675 1680
 Val Ser Gln Thr Ser Leu Leu Glu Ala Lys Lys Trp Leu Arg Glu Gly
 1685 1690 1695
 Ile Phe Asp Gly Gln Pro Glu Arg Ile Asn Thr Ala Asp Tyr Val Gly
 1700 1705 1710
 Asn Tyr Leu Tyr Glu Asn Asn Ser Asn Ser Thr Ile Ala Glu Asn Asp
 1715 1720 1725
 Lys Asn His Leu Ser Glu Lys Gln Asp Thr Tyr Leu Ser Asn Ser Ser
 1730 1735 1740
 Met Ser Asn Ser Tyr Ser Tyr His Ser Asp Glu Val Tyr Asn Asp Ser
 1745 1750 1755 1760
 Gly Tyr Leu Ser Lys Asn Lys Leu Asp Ser Gly Ile Glu Pro Val Leu
 1765 1770 1775
 Lys Asn Val Glu Asp Gln Lys Asn Thr Ser Phe Ser Lys Val Ile Ser
 1780 1785 1790
 Asn Val Lys Asp Ala Asn Ala Tyr Pro Gln Thr Val Asn Glu Asp Ile
 1795 1800 1805
 Cys Val Glu Glu Leu Val Thr Ser Ser Pro Cys Lys Asn Lys Asn
 1810 1815 1820
 Ala Ala Ile Lys Leu Ser Ile Ser Asn Ser Asn Asn Phe Glu Val Gly
 1825 1830 1835 1840
 Pro Pro Ala Phe Arg Ile Ala Ser Gly Lys Ile Val Cys Val Ser His
 1845 1850 1855
 Glu Thr Ile Lys Lys Val Lys Asp Ile Phe Thr Asp Ser Phe Ser Lys
 1860 1865 1870
 Val Ile Lys Glu Asn Asn Glu Asn Lys Ser Lys Ile Cys Gln Thr Lys
 1875 1880 1885
 Ile Met Ala Gly Cys Tyr Glu Ala Leu Asp Asp Ser Glu Asp Ile Leu
 1890 1895 1900
 His Asn Ser Leu Asp Asn Asp Glu Cys Ser Thr His Ser His Lys Val
 1905 1910 1915 1920
 Phe Ala Asp Ile Gln Ser Glu Glu Ile Leu Gln His Asn Gln Asn Met
 1925 1930 1935
 Ser Gly Leu Glu Lys Val Ser Lys Ile Ser Pro Cys Asp Val Ser Leu
 1940 1945 1950
 Glu Thr Ser Asp Ile Cys Lys Cys Ser Ile Gly Lys Leu His Lys Ser
 1955 1960 1965
 Val Ser Ser Ala Asn Thr Cys Gly Ile Phe Ser Thr Ala Ser Gly Lys
 1970 1975 1980
 Ser Val Gln Val Ser Asp Ala Ser Leu Gln Asn Ala Arg Gln Val Phe
 1985 1990 1995 2000
 Ser Glu Ile Glu Asp Ser Thr Lys Gln Val Phe Ser Lys Val Leu Phe
 2005 2010 2015
 Lys Ser Asn Glu His Ser Asp Gln Leu Thr Arg Glu Glu Asn Thr Ala
 2020 2025 2030
 Ile Arg Thr Pro Glu His Leu Ile Ser Gln Lys Gly Phe Ser Tyr Asn
 2035 2040 2045
 Val Val Asn Ser Ser Ala Phe Ser Gly Phe Ser Thr Ala Ser Gly Lys
 2050 2055 2060
 Gln Val Ser Ile Leu Glu Ser Ser Leu His Lys Val Lys Gly Val Leu
 2065 2070 2075 2080
 Glu Glu Phe Asp Leu Ile Arg Thr Glu His Ser Leu His Tyr Ser Pro
 2085 2090 2095
 Thr Ser Arg Gln Asn Val Ser Lys Ile Leu Pro Arg Val Asp Lys Arg
 2100 2105 2110
 Asn Pro Glu His Cys Val Asn Ser Glu Met Glu Lys Thr Cys Ser Lys
 2115 2120 2125
 Glu Phe Lys Leu Ser Asn Asn Leu Asn Val Glu Gly Gly Ser Ser Glu
 2130 2135 2140
 Asn Asn His Ser Ile Lys Val Ser Pro Tyr Leu Ser Gln Phe Gln Gln

2145 2150 2155 2160
 Asp Lys Gln Gln Leu Val Leu Gly Thr Lys Val Ser Leu Val Glu Asn
 2165 2170 2175
 Ile His Val Leu Gly Lys Glu Gln Ala Ser Pro Lys Asn Val Lys Met
 2180 2185 2190
 Glu Ile Gly Lys Thr Glu Thr Phe Ser Asp Val Pro Val Lys Thr Asn
 2195 2200 2205
 Ile Glu Val Cys Ser Thr Tyr Ser Lys Asp Ser Glu Asn Tyr Phe Glu
 2210 2215 2220
 Thr Glu Ala Val Glu Ile Ala Lys Ala Phe Met Glu Asp Asp Glu Leu
 2225 2230 2235 2240
 Thr Asp Ser Lys Leu Pro Ser His Ala Thr His Ser Leu Phe Thr Cys
 2245 2250 2255
 Pro Glu Asn Glu Glu Met Val Leu Ser Asn Ser Arg Ile Gly Lys Arg
 2260 2265 2270
 Arg Gly Glu Pro Leu Ile Leu Val Gly Glu Pro Ser Ile Lys Arg Asn
 2275 2280 2285
 Leu Leu Asn Glu Phe Asp Arg Ile Ile Glu Asn Gln Glu Lys Ser Leu
 2290 2295 2300
 Lys Ala Ser Lys Ser Thr Pro Asp Gly Thr Ile Lys Asp Arg Arg Leu
 2305 2310 2315 2320
 Phe Met His His Val Ser Leu Glu Pro Ile Thr Cys Val Pro Phe Arg
 2325 2330 2335
 Thr Thr Lys Glu Arg Gln Glu Ile Gln Asn Pro Asn Phe Thr Ala Pro
 2340 2345 2350
 Gly Gln Glu Phe Leu Ser Lys Ser His Leu Tyr Glu His Leu Thr Leu
 2355 2360 2365
 Glu Lys Ser Ser Ser Asn Leu Ala Val Ser Gly His Pro Phe Tyr Gln
 2370 2375 2380
 Val Ser Ala Thr Arg Asn Glu Lys Met Arg His Leu Ile Thr Thr Gly
 2385 2390 2395 2400
 Arg Pro Thr Lys Val Phe Val Pro Pro Phe Lys Thr Lys Ser His Phe
 2405 2410 2415
 His Arg Val Glu Gln Cys Val Arg Asn Ile Asn Leu Glu Glu Asn Arg
 2420 2425 2430
 Gln Lys Gln Asn Ile Asp Gly His Gly Ser Asp Asp Ser Lys Asn Lys
 2435 2440 2445
 Ile Asn Asp Asn Glu Ile His Gln Phe Asn Lys Asn Asn Ser Asn Gln
 2450 2455 2460
 Ala Ala Ala Val Thr Phe Thr Lys Cys Glu Glu Pro Leu Asp Leu
 2465 2470 2475 2480
 Ile Thr Ser Leu Gln Asn Ala Arg Asp Ile Gln Asp Met Arg Ile Lys
 2485 2490 2495
 Lys Lys Gln Arg Gln Arg Val Phe Pro Gln Pro Gly Ser Leu Tyr Leu
 2500 2505 2510
 Ala Lys Thr Ser Thr Leu Pro Arg Ile Ser Leu Lys Ala Ala Val Gly
 2515 2520 2525
 Gly Gln Val Pro Ser Ala Cys Ser His Lys Gln Leu Tyr Thr Tyr Gly
 2530 2535 2540
 Val Ser Lys His Cys Ile Lys Ile Asn Ser Lys Asn Ala Glu Ser Phe
 2545 2550 2555 2560
 Gln Phe His Thr Glu Asp Tyr Phe Gly Lys Glu Ser Leu Trp Thr Gly
 2565 2570 2575
 Lys Gly Ile Gln Leu Ala Asp Gly Gly Trp Leu Ile Pro Ser Asn Asp
 2580 2585 2590
 Gly Lys Ala Gly Lys Glu Glu Phe Tyr Arg Ala Leu Cys Asp Thr Pro
 2595 2600 2605
 Gly Val Asp Pro Lys Leu Ile Ser Arg Ile Trp Val Tyr Asn His Tyr
 2610 2615 2620
 Arg Trp Ile Ile Trp Lys Leu Ala Ala Met Glu Cys Ala Phe Pro Lys
 2625 2630 2635 2640

Glu Phe Ala Asn Arg Cys Leu Ser Pro Glu Arg Val Leu Leu Gln Leu
 2645 2650 2655
 Lys Tyr Arg Tyr Asp Thr Glu Ile Asp Arg Ser Arg Arg Ser Ala Ile
 2660 2665 2670
 Lys Lys Ile Met Glu Arg Asp Asp Thr Ala Ala Lys Thr Leu Val Leu
 2675 2680 2685
 Cys Val Ser Asp Ile Ile Ser Leu Ser Ala Asn Ile Ser Glu Thr Ser
 2690 2695 2700
 Ser Asn Lys Thr Ser Ser Ala Asp Thr Gln Lys Val Ala Ile Ile Glu
 2705 2710 2715 2720
 Leu Thr Asp Gly Trp Tyr Ala Val Lys Ala Gln Leu Asp Pro Pro Leu
 2725 2730 2735
 Leu Ala Val Leu Lys Asn Gly Arg Leu Thr Val Gly Gln Lys Ile Ile
 2740 2745 2750
 Leu His Gly Ala Glu Leu Val Gly Ser Pro Asp Ala Cys Thr Pro Leu
 2755 2760 2765
 Glu Ala Pro Glu Ser Leu Met Leu Lys Ile Ser Ala Asn Ser Thr Arg
 2770 2775 2780
 Pro Ala Arg Trp Tyr Thr Lys Leu Gly Phe Phe Pro Asp Pro Arg Pro
 2785 2790 2795 2800
 Phe Pro Leu Pro Leu Ser Ser Leu Phe Ser Asp Gly Gly Asn Val Gly
 2805 2810 2815
 Cys Val Asp Val Ile Ile Gln Arg Ala Tyr Pro Ile Gln Trp Met Glu
 2820 2825 2830
 Lys Thr Ser Ser Gly Leu Tyr Ile Phe Arg Asn Glu Arg Glu Glu Glu
 2835 2840 2845
 Lys Glu Ala Ala Lys Tyr Val Glu Ala Gln Gln Lys Arg Leu Glu Ala
 2850 2855 2860
 Leu Phe Thr Lys Ile Gln Glu Glu Phe Glu Glu His Glu Glu Asn Thr
 2865 2870 2875 2880
 Thr Lys Pro Tyr Leu Pro Ser Arg Ala Leu Thr Arg Gln Gln Val Arg
 2885 2890 2895
 Ala Leu Gln Asp Gly Ala Glu Leu Tyr Glu Ala Val Lys Asn Ala Ala
 2900 2905 2910
 Asp Pro Ala Tyr Leu Glu Gly Tyr Phe Ser Glu Glu Gln Leu Arg Ala
 2915 2920 2925
 Leu Asn Asn His Arg Gln Met Leu Asn Asp Lys Lys Gln Ala Gln Ile
 2930 2935 2940
 Gln Leu Glu Ile Arg Lys Ala Met Glu Ser Ala Glu Gln Lys Glu Gln
 2945 2950 2955 2960
 Gly Leu Ser Arg Asp Val Thr Thr Val Trp Lys Leu Arg Ile Val Ser
 2965 2970 2975
 Tyr Ser Lys Lys Glu Lys Asp Ser Val Ile Leu Ser Ile Trp Arg Pro
 2980 2985 2990
 Ser Ser Asp Leu Tyr Ser Leu Leu Thr Glu Gly Lys Arg Tyr Arg Ile
 2995 3000 3005
 Tyr His Leu Ala Thr Ser Lys Ser Lys Ser Lys Ser Glu Arg Ala Asn
 3010 3015 3020
 Ile Gln Leu Ala Ala Thr Lys Lys Thr Gln Tyr Gln Gln Leu Pro Val
 3025 3030 3035 3040
 Ser Asp Glu Ile Leu Phe Gln Ile Tyr Gln Pro Arg Glu Pro Leu His
 3045 3050 3055
 Phe Ser Lys Phe Leu Asp Pro Asp Phe Gln Pro Ser Cys Ser Glu Val
 3060 3065 3070
 Asp Leu Ile Gly Phe Val Val Ser Val Val Lys Lys Thr Gly Leu Ala
 3075 3080 3085
 Pro Phe Val Tyr Leu Ser Asp Glu Cys Tyr Asn Leu Leu Ala Ile Lys
 3090 3095 3100
 Phe Trp Ile Asp Leu Asn Glu Asp Ile Ile Lys Pro His Met Leu Ile
 3105 3110 3115 3120
 Ala Ala Ser Asn Leu Gln Trp Arg Pro Glu Ser Lys Ser Gly Leu Leu

| | | |
|---|------|------|
| 3125 | 3130 | 3135 |
| Thr Leu Phe Ala Gly Asp Phe Ser Val Phe Ser Ala Ser Pro Lys Glu | | |
| 3140 | 3145 | 3150 |
| Gly His Phe Gln Glu Thr Phe Asn Lys Met Lys Asn Thr Val Glu Asn | | |
| 3155 | 3160 | 3165 |
| Ile Asp Ile Leu Cys Asn Glu Ala Glu Asn Lys Leu Met His Ile Leu | | |
| 3170 | 3175 | 3180 |
| His Ala Asn Asp Pro Lys Trp Ser Thr Pro Thr Lys Asp Cys Thr Ser | | |
| 3185 | 3190 | 3195 |
| Gly Pro Tyr Thr Ala Gln Ile Ile Pro Gly Thr Gly Asn Lys Leu Leu | | |
| 3205 | 3210 | 3215 |
| Met Ser Ser Pro Asn Cys Glu Ile Tyr Tyr Gln Ser Pro Leu Ser Leu | | |
| 3220 | 3225 | 3230 |
| Cys Met Ala Lys Arg Lys Ser Val Ser Thr Pro Val Ser Ala Gln Met | | |
| 3235 | 3240 | 3245 |
| Thr Ser Lys Ser Cys Lys Gly Glu Lys Glu Ile Asp Asp Gln Lys Asn | | |
| 3250 | 3255 | 3260 |
| Cys Lys Lys Arg Arg Ala Leu Asp Phe Leu Ser Arg Leu Pro Leu Pro | | |
| 3265 | 3270 | 3275 |
| Pro Pro Val Ser Pro Ile Cys Thr Phe Val Ser Pro Ala Ala Gln Lys | | |
| 3285 | 3290 | 3295 |
| Ala Phe Gln Pro Pro Arg Ser Cys Gly Thr Lys Tyr Glu Thr Pro Ile | | |
| 3300 | 3305 | 3310 |
| Lys Lys Lys Glu Leu Asn Ser Pro Gln Met Thr Pro Phe Lys Lys Phe | | |
| 3315 | 3320 | 3325 |
| Asn Glu Ile Ser Leu Leu Glu Ser Asn Ser Ile Ala Asp Glu Glu Leu | | |
| 3330 | 3335 | 3340 |
| Ala Leu Ile Asn Thr Gln Ala Leu Leu Ser Gly Ser Thr Gly Glu Lys | | |
| 3345 | 3350 | 3355 |
| Gln Phe Ile Ser Val Ser Glu Ser Thr Arg Thr Ala Pro Thr Ser Ser | | |
| 3365 | 3370 | 3375 |
| Glu Asp Tyr Leu Arg Leu Lys Arg Arg Cys Thr Thr Ser Leu Ile Lys | | |
| 3380 | 3385 | 3390 |
| Glu Gln Glu Ser Ser Gln Ala Ser Thr Glu Glu Cys Glu Lys Asn Lys | | |
| 3395 | 3400 | 3405 |
| Gln Asp Thr Ile Thr Thr Lys Lys Tyr Ile | | |
| 3410 | 3415 | |

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10485 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
(B) LOCATION: 229...10482
(D) OTHER INFORMATION: BRCA2 (OMIM3)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GGTGGCGCGA GCTTCTGAAA CTAGGCGGCA GAGGC GGAGC CGCTGTGGCA CTGCTGCGCC
 TCTGCTGCGC CTCGGGTGTC TTTTGC GGCG GTGGGTCGCC GCCGGGAGAA GCGTGAGGGG
 ACAGATTGTC GACCGGGCGCG GTTTTGTCA GCTTACTCCG GCCAAAAAAAG AACTGCACCT
 CTGGAGCGGA CTTATTTACC AAGCATTGGA GGAATATCGT AGGTAAAAA ATG CCT ATT
Met Pro Ile

GGA TCC AAA GAG AGG CCA ACA TTT TTT GAA ATT TTT AAG ACA CGC TGC 285
 Gly Ser Lys Glu Arg Pro Thr Phe Phe Glu Ile Phe Lys Thr Arg Cys
 5 10 15

AAC AAA GCA GAT TTA GGA CCA ATA AGT CTT AAT TGG TTT GAA GAA CTT 333
 Asn Lys Ala Asp Leu Gly Pro Ile Ser Leu Asn Trp Phe Glu Glu Leu
 20 25 30 35

TCT TCA GAA GCT CCA CCC TAT AAT TCT GAA CCT GCA GAA GAA TCT GAA 381
 Ser Ser Glu Ala Pro Pro Tyr Asn Ser Glu Pro Ala Glu Glu Ser Glu
 40 45 50

CAT AAA AAC AAC AAT TAC GAA CCA AAC CTA TTT AAA ACT CCA CAA AGG 429
 His Lys Asn Asn Asn Tyr Glu Pro Asn Leu Phe Lys Thr Pro Gln Arg
 55 60 65

AAA CCA TCT TAT AAT CAG CTG GCT TCA ACT CCA ATA ATA TTC AAA GAG 477
 Lys Pro Ser Tyr Asn Gln Leu Ala Ser Thr Pro Ile Ile Phe Lys Glu
 70 75 80

CAA GGG CTG ACT CTG CCG CTG TAC CAA TCT CCT GTA AAA GAA TTA GAT 525
 Gln Gly Leu Thr Leu Pro Leu Tyr Gln Ser Pro Val Lys Glu Leu Asp
 85 90 95

AAA TTC AAA TTA GAC TTA GGA AGG AAT GTT CCC AAT AGT AGA CAT AAA 573
 Lys Phe Lys Leu Asp Leu Gly Arg Asn Val Pro Asn Ser Arg His Lys
 100 105 110 115

AGT CTT CGC ACA GTG AAA ACT AAA ATG GAT CAA GCA GAT GAT GTT TCC 621
 Ser Leu Arg Thr Val Lys Thr Lys Met Asp Gln Ala Asp Asp Val Ser
 120 125 130

TGT CCA CTT CTA AAT TCT TGT CTT AGT GAA AGT CCT GTT GTT CTA CAA 669
 Cys Pro Leu Leu Asn Ser Cys Leu Ser Glu Ser Pro Val Val Leu Gln
 135 140 145

TGT ACA CAT GTA ACA CCA CAA AGA GAT AAG TCA GTG GTA TGT GGG AGT 717
 Cys Thr His Val Thr Pro Gln Arg Asp Lys Ser Val Val Cys Gly Ser
 150 155 160

TTG TTT CAT ACA CCA AAG TTT GTG AAG GGT CGT CAG ACA CCA AAA CAT 765
 Leu Phe His Thr Pro Lys Phe Val Lys Gly Arg Gln Thr Pro Lys His
 165 170 175

ATT TCT GAA AGT CTA GGA GCT GAG GTG GAT CCT GAT ATG TCT TGG TCA 813
 Ile Ser Glu Ser Leu Gly Ala Glu Val Asp Pro Asp Met Ser Trp Ser
 180 185 190 195

AGT TCT TTA GCT ACA CCA CCC ACC CTT AGT TCT ACT GTG CTC ATA GTC 861
 Ser Ser Leu Ala Thr Pro Pro Thr Leu Ser Ser Thr Val Leu Ile Val
 200 205 210

AGA AAT GAA GAA GCA TCT GAA ACT GTA TTT CCT CAT GAT ACT ACT GCT 909
 Arg Asn Glu Glu Ala Ser Glu Thr Val Phe Pro His Asp Thr Thr Ala
 215 220 225

AAT GTG AAA AGC TAT TTT TCC AAT CAT GAT GAA AGT CTG AAG AAA AAT 957
 Asn Val Lys Ser Tyr Phe Ser Asn His Asp Glu Ser Leu Lys Lys Asn
 230 235 240

GAT AGA TTT ATC GCT TCT GTG ACA GAC AGT GAA AAC ACA AAT CAA AGA 1005
 Asp Arg Phe Ile Ala Ser Val Thr Asp Ser Glu Asn Thr Asn Gln Arg
 245 250 255

GAA GCT GCA AGT CAT GGA TTT GGA AAA ACA TCA GGG AAT TCA TTT AAA 1053
 Glu Ala Ala Ser His Gly Phe Gly Lys Thr Ser Gly Asn Ser Phe Lys
 260 265 270 275

GTA AAT AGC TGC AAA GAC CAC ATT GGA AAG TCA ATG CCA CAT GTC CTA 1101
 Val Asn Ser Cys Lys Asp His Ile Gly Lys Ser Met Pro His Val Leu
 280 285 290

GAA GAT GAA GTA TAT GAA ACA GTT GTA GAT ACC TCT GAA GAA GAT AGT 1149
 Glu Asp Glu Val Tyr Glu Thr Val Val Asp Thr Ser Glu Glu Asp Ser
 295 300 305

TTT TCA TTA TGT TTT TCT AAA TGT AGA ACA AAA AAT CTA CAA AAA GTA 1197
 Phe Ser Leu Cys Phe Ser Lys Cys Arg Thr Lys Asn Leu Gln Lys Val
 310 315 320

AGA ACT AGC AAG ACT AGG AAA AAA ATT TTC CAT GAA GCA AAC GCT GAT 1245
 Arg Thr Ser Lys Thr Arg Lys Lys Ile Phe His Glu Ala Asn Ala Asp
 325 330 335

GAA TGT GAA AAA TCT AAA AAC CAA GTG AAA GAA AAA TAC TCA TTT GTA 1293
 Glu Cys Glu Lys Ser Lys Asn Gln Val Lys Glu Lys Tyr Ser Phe Val
 340 345 350 355

TCT GAA GTG GAA CCA AAT GAT ACT GAT CCA TTA GAT TCA AAT GTA GCA 1341
 Ser Glu Val Glu Pro Asn Asp Thr Asp Pro Leu Asp Ser Asn Val Ala
 360 365 370

AAT CAG AAG CCC TTT GAG AGT GGA AGT GAC AAA ATC TCC AAG GAA GTT 1389
 Asn Gln Lys Pro Phe Glu Ser Gly Ser Asp Lys Ile Ser Lys Glu Val
 375 380 385

GTA CCG TCT TTG GCC TGT GAA TGG TCT CAA CTA ACC CTT TCA GGT CTA 1437
 Val Pro Ser Leu Ala Cys Glu Trp Ser Gln Leu Thr Leu Ser Gly Leu
 390 395 400

AAT GGA GCC CAG ATG GAG AAA ATA CCC CTA TTG CAT ATT TCT TCA TGT 1485
 Asn Gly Ala Gln Met Glu Lys Ile Pro Leu Leu His Ile Ser Ser Cys
 405 410 415

GAC CAA AAT ATT TCA GAA AAA GAC CTA TTA GAC ACA GAG AAC AAA AGA 1533
 Asp Gln Asn Ile Ser Glu Lys Asp Leu Leu Asp Thr Glu Asn Lys Arg
 420 425 430 435

AAG AAA GAT TTT CTT ACT TCA GAG AAT TCT TTG CCA CGT ATT TCT AGC 1581
 Lys Lys Asp Phe Leu Thr Ser Glu Asn Ser Leu Pro Arg Ile Ser Ser
 440 445 450

CTA CCA AAA TCA GAG AAG CCA TTA AAT GAG GAA ACA GTG GTA AAT AAG 1629
 Leu Pro Lys Ser Glu Lys Pro Leu Asn Glu Glu Thr Val Val Asn Lys
 455 460 465

AGA GAT GAA GAG CAG CAT CTT GAA TCT CAT ACA GAC TGC ATT CTT GCA 1677
 Arg Asp Glu Glu Gln His Leu Glu Ser His Thr Asp Cys Ile Leu Ala
 470 475 480

| | |
|---|------|
| GTA AAG CAG GCA ATA TCT GGA ACT TCT CCA GTG GCT TCT TCA TTT CAG Val Lys Gln Ala Ile Ser Gly Thr Ser Pro Val Ala Ser Ser Phe Gln 485 490 495 | 1725 |
| GGT ATC AAA AAG TCT ATA TTC AGA ATA AGA GAA TCA CCT AAA GAG ACT Gly Ile Lys Lys Ser Ile Phe Arg Ile Arg Glu Ser Pro Lys Glu Thr 500 505 510 515 | 1773 |
| TTC AAT GCA AGT TTT TCA GGT CAT ATG ACT GAT CCA AAC TTT AAA AAA Phe Asn Ala Ser Phe Ser Gly His Met Thr Asp Pro Asn Phe Lys Lys 520 525 530 | 1821 |
| GAA ACT GAA GCC TCT GAA AGT GGA CTG GAA ATA CAT ACT GTT TGC TCA Glu Thr Glu Ala Ser Glu Ser Gly Leu Glu Ile His Thr Val Cys Ser 535 540 545 | 1869 |
| CAG AAG GAG GAC TCC TTA TGT CCA AAT TTA ATT GAT AAT GGA AGC TGG Gln Lys Glu Asp Ser Leu Cys Pro Asn Leu Ile Asp Asn Gly Ser Trp 550 555 560 | 1917 |
| CCA GCC ACC ACC ACA CAG AAT TCT GTA GCT TTG AAG AAT GCA GGT TTA Pro Ala Thr Thr Gln Asn Ser Val Ala Leu Lys Asn Ala Gly Leu 565 570 575 | 1965 |
| ATA TCC ACT TTG AAA AAG AAA ACA AAT AAG TTT ATT TAT GCT ATA CAT Ile Ser Thr Leu Lys Lys Thr Asn Lys Phe Ile Tyr Ala Ile His 580 585 590 595 | 2013 |
| GAT GAA ACA TCT TAT AAA GGA AAA AAA ATA CCG AAA GAC CAA AAA TCA Asp Glu Thr Ser Tyr Lys Gly Lys Ile Pro Lys Asp Gln Lys Ser 600 605 610 | 2061 |
| GAA CTA ATT AAC TGT TCA GCC CAG TTT GAA GCA AAT GCT TTT GAA GCA Glu Leu Ile Asn Cys Ser Ala Gln Phe Glu Ala Asn Ala Phe Glu Ala 615 620 625 | 2109 |
| CCA CTT ACA TTT GCA AAT GCT GAT TCA GGT TTA TTG CAT TCT TCT GTG Pro Leu Thr Phe Ala Asn Ala Asp Ser Gly Leu Leu His Ser Ser Val 630 635 640 | 2157 |
| AAA AGA AGC TGT TCA CAG AAT GAT TCT GAA GAA CCA ACT TTG TCC TTA Lys Arg Ser Cys Ser Gln Asn Asp Ser Glu Glu Pro Thr Leu Ser Leu 645 650 655 | 2205 |
| ACT AGC TCT TTT GGG ACA ATT CTG AGG AAA TGT TCT AGA AAT GAA ACA Thr Ser Ser Phe Gly Thr Ile Leu Arg Lys Cys Ser Arg Asn Glu Thr 660 665 670 675 | 2253 |
| TGT TCT AAT AAT ACA GTA ATC TCT CAG GAT CTT GAT TAT AAA GAA GCA Cys Ser Asn Asn Thr Val Ile Ser Gln Asp Leu Asp Tyr Lys Glu Ala 680 685 690 | 2301 |
| AAA TGT AAT AAG GAA AAA CTA CAG TTA TTT ATT ACC CCA GAA GCT GAT Lys Cys Asn Lys Glu Lys Leu Gln Phe Ile Thr Pro Glu Ala Asp 695 700 705 | 2349 |
| TCT CTG TCA TGC CTG CAG GAA GGA CAG TGT GAA AAT GAT CCA AAA AGC Ser Leu Ser Cys Leu Gln Glu Gly Gln Cys Glu Asn Asp Pro Lys Ser 710 715 720 | 2397 |
| AAA AAA GTT TCA GAT ATA AAA GAA GAG GTC TTG GCT GCA GCA TGT CAC | 2445 |

Lys Lys Val Ser Asp Ile Lys Glu Glu Val Leu Ala Ala Ala Cys His
 725 730 735

CCA GTA CAA CAC TCA AAA GTG GAA TAC AGT GAT ACT GAC TTT CAA TCC 2493
 Pro Val Gln His Ser Lys Val Glu Tyr Ser Asp Thr Asp Phe Gln Ser
 740 745 750 755

CAG AAA AGT CTT TTA TAT GAT CAT GAA AAT GCC AGC ACT CTT ATT TTA 2541
 Gln Lys Ser Leu Leu Tyr Asp His Glu Asn Ala Ser Thr Leu Ile Leu
 760 765 770

ACT CCT ACT TCC AAG GAT GTT CTG TCA AAC CTA GTC ATG ATT TCT AGA 2589
 Thr Pro Thr Ser Lys Asp Val Leu Ser Asn Leu Val Met Ile Ser Arg
 775 780 785

GGC AAA GAA TCA TAC AAA ATG TCA GAC AAG CTC AAA GGT AAC AAT TAT 2637
 Gly Lys Glu Ser Tyr Lys Met Ser Asp Lys Leu Lys Gly Asn Asn Tyr
 790 795 800

GAA TCT GAT GTT GAA TTA ACC AAA AAT ATT CCC ATG GAA AAG AAT CAA 2685
 Glu Ser Asp Val Glu Leu Thr Lys Asn Ile Pro Met Glu Lys Asn Gln
 805 810 815

GAT GTA TGT GCT TTA AAT GAA AAT TAT AAA AAC GTT GAG CTG TTG CCA 2733
 Asp Val Cys Ala Leu Asn Glu Asn Tyr Lys Asn Val Glu Leu Leu Pro
 820 825 830 835

CCT GAA AAA TAC ATG AGA GTA GCA TCA CCT TCA AGA AAG GTA CAA TTC 2781
 Pro Glu Lys Tyr Met Arg Val Ala Ser Pro Ser Arg Lys Val Gln Phe
 840 845 850

AAC CAA AAC ACA AAT CTA AGA GTA ATC CAA AAA AAT CAA GAA GAA ACT 2829
 Asn Gln Asn Thr Asn Leu Arg Val Ile Gln Lys Asn Gln Glu Thr
 855 860 865

ACT TCA ATT TCA AAA ATA ACT GTC AAT CCA GAC TCT GAA GAA CTT TTC 2877
 Thr Ser Ile Ser Lys Ile Thr Val Asn Pro Asp Ser Glu Glu Leu Phe
 870 875 880

TCA GAC AAT GAG AAT AAT TTT GTC TTC CAA GTA GCT AAT GAA AGG AAT 2925
 Ser Asp Asn Glu Asn Asn Phe Val Phe Gln Val Ala Asn Glu Arg Asn
 885 890 895

AAT CTT GCT TTA GGA AAT ACT AAG GAA CTT CAT GAA ACA GAC TTG ACT 2973
 Asn Leu Ala Leu Gly Asn Thr Lys Glu Leu His Glu Thr Asp Leu Thr
 900 905 910 915

TGT GTA AAC GAA CCC ATT TTC AAG AAC TCT ACC ATG GTT TTA TAT GGA 3021
 Cys Val Asn Glu Pro Ile Phe Lys Asn Ser Thr Met Val Leu Tyr Gly
 920 925 930

GAC ACA GGT GAT AAA CAA GCA ACC CAA GTG TCA ATT AAA AAA GAT TTG 3069
 Asp Thr Gly Asp Lys Gln Ala Thr Gln Val Ser Ile Lys Lys Asp Leu
 935 940 945

GTT TAT GTT CTT GCA GAG GAG AAC AAA AAT AGT GTA AAG CAG CAT ATA 3117
 Val Tyr Val Leu Ala Glu Glu Asn Lys Asn Ser Val Lys Gln His Ile
 950 955 960

AAA ATG ACT CTA GGT CAA GAT TTA AAA TCG GAC ATC TCC TTG AAT ATA 3165
 Lys Met Thr Leu Gly Gln Asp Leu Lys Ser Asp Ile Ser Leu Asn Ile

965 970 975 3213
 GAT AAA ATA CCA GAA AAA AAT AAT GAT TAC ATG GAC AAA TGG GCA GGA
 Asp Lys Ile Pro Glu Lys Asn Asn Asp Tyr Met Asp Lys Trp Ala Gly
 980 985 990 995

 CTC TTA GGT CCA ATT TCA AAT CAC AGT TTT GGA GGT AGC TTC AGA ACA 3261
 Leu Leu Gly Pro Ile Ser Asn His Ser Phe Gly Gly Ser Phe Arg Thr
 1000 1005 1010

 GCT TCA AAT AAG GAA ATC AAG CTC TCT GAA CAT AAC ATT AAG AAG AGC 3309
 Ala Ser Asn Lys Glu Ile Lys Leu Ser Glu His Asn Ile Lys Lys Ser
 1015 1020 1025

 AAA ATG TTC TTC AAA GAT ATT GAA GAA CAA TAT CCT ACT AGT TTA GCT 3357
 Lys Met Phe Phe Lys Asp Ile Glu Glu Gln Tyr Pro Thr Ser Leu Ala
 1030 1035 1040

 TGT GTT GAA ATT GTA AAT ACC TTG GCA TTA GAT AAT CAA AAG AAA CTG 3405
 Cys Val Glu Ile Val Asn Thr Leu Ala Leu Asp Asn Gln Lys Lys Leu
 1045 1050 1055

 AGC AAG CCT CAG TCA ATT AAT ACT GTA TCT GCA CAT TTA CAG AGT AGT 3453
 Ser Lys Pro Gln Ser Ile Asn Thr Val Ser Ala His Leu Gln Ser Ser
 1060 1065 1070 1075

 GTA GTT GTT TCT GAT TGT AAA AAT AGT CAT ATA ACC CCT CAG ATG TTA 3501
 Val Val Val Ser Asp Cys Lys Asn Ser His Ile Thr Pro Gln Met Leu
 1080 1085 1090

 TTT TCC AAG CAG GAT TTT AAT TCA AAC CAT AAT TTA ACA CCT AGC CAA 3549
 Phe Ser Lys Gln Asp Phe Asn Ser Asn His Asn Leu Thr Pro Ser Gln
 1095 1100 1105

 AAG GCA GAA ATT ACA GAA CTT TCT ACT ATA TTA GAA GAA TCA GGA AGT 3597
 Lys Ala Glu Ile Thr Glu Leu Ser Thr Ile Leu Glu Ser Gly Ser
 1110 1115 1120

 CAG TTT GAA TTT ACT CAG TTT AGA AAG CCA AGC TAC ATA TTG CAG AAG 3645
 Gln Phe Glu Phe Thr Gln Phe Arg Lys Pro Ser Tyr Ile Leu Gln Lys
 1125 1130 1135

 AGT ACA TTT GAA GTG CCT GAA AAC CAG ATG ACT ATC TTA AAG ACC ACT 3693
 Ser Thr Phe Glu Val Pro Glu Asn Gln Met Thr Ile Leu Lys Thr Thr
 1140 1145 1150 1155

 TCT GAG GAA TGC AGA GAT GCT GAT CTT CAT GTC ATA ATG AAT GCC CCA 3741
 Ser Glu Glu Cys Arg Asp Ala Asp Leu His Val Ile Met Asn Ala Pro
 1160 1165 1170

 TCG ATT GGT CAG GTA GAC AGC AGC AAG CAA TTT GAA GGT ACA GTT GAA 3789
 Ser Ile Gly Gln Val Asp Ser Ser Lys Gln Phe Glu Gly Thr Val Glu
 1175 1180 1185

 ATT AAA CGG AAG TTT GCT GGC CTG TTG AAA AAT GAC TGT AAC AAA AGT 3837
 Ile Lys Arg Lys Phe Ala Gly Leu Leu Lys Asn Asp Cys Asn Lys Ser
 1190 1195 1200

 GCT TCT GGT TAT TTA ACA GAT GAA AAT GAA GTG GGG TTT AGG GGC TTT 3885
 Ala Ser Gly Tyr Leu Thr Asp Glu Asn Glu Val Gly Phe Arg Gly Phe
 1205 1210 1215

TAT TCT GCT CAT GGC ACA AAA CTG AAT GTT TCT ACT GAA GCT CTG CAA 3933
 Tyr Ser Ala His Gly Thr Lys Leu Asn Val Ser Thr Glu Ala Leu Gln
 1220 1225 1230 1235

AAA GCT GTG AAA CTG TTT AGT GAT ATT GAG AAT ATT AGT GAG GAA ACT 3981
 Lys Ala Val Lys Leu Phe Ser Asp Ile Glu Asn Ile Ser Glu Glu Thr
 1240 1245 1250

TCT GCA GAG GTA CAT CCA ATA AGT TTA TCT TCA AGT AAA TGT CAT GAT 4029
 Ser Ala Glu Val His Pro Ile Ser Leu Ser Ser Ser Lys Cys His Asp
 1255 1260 1265

TCT GTT GTT TCA ATG TTT AAG ATA GAA AAT CAT AAT GAT AAA ACT GTA 4077
 Ser Val Val Ser Met Phe Lys Ile Glu Asn His Asn Asp Lys Thr Val
 1270 1275 1280

AGT GAA AAA AAT AAT AAA TGC CAA CTG ATA TTA CAA AAT AAT ATT GAA 4125
 Ser Glu Lys Asn Asn Lys Cys Gln Leu Ile Leu Gln Asn Asn Ile Glu
 1285 1290 1295

ATG ACT ACT GGC ACT TTT GTT GAA GAA ATT ACT GAA AAT TAC AAG AGA 4173
 Met Thr Thr Gly Thr Phe Val Glu Glu Ile Thr Glu Asn Tyr Lys Arg
 1300 1305 1310 1315

AAT ACT GAA AAT GAA GAT AAC AAA TAT ACT GCT GCC AGT AGA AAT TCT 4221
 Asn Thr Glu Asn Glu Asp Asn Lys Tyr Thr Ala Ala Ser Arg Asn Ser
 1320 1325 1330

CAT AAC TTA GAA TTT GAT GGC AGT GAT TCA AGT AAA AAT GAT ACT GTT 4269
 His Asn Leu Glu Phe Asp Gly Ser Asp Ser Ser Lys Asn Asp Thr Val
 1335 1340 1345

TGT ATT CAT AAA GAT GAA ACG GAC TTG CTA TTT ACT GAT CAG CAC AAC 4317
 Cys Ile His Lys Asp Glu Thr Asp Leu Leu Phe Thr Asp Gln His Asn
 1350 1355 1360

ATA TGT CTT AAA TTA TCT GGC CAG TTT ATG AAG GAG GGA AAC ACT CAG 4365
 Ile Cys Leu Lys Leu Ser Gly Gln Phe Met Lys Glu Gly Asn Thr Gln
 1365 1370 1375

ATT AAA GAA GAT TTG TCA GAT TTA ACT TTT TTG GAA GTT GCG AAA GCT 4413
 Ile Lys Glu Asp Leu Ser Asp Leu Thr Phe Leu Glu Val Ala Lys Ala
 1380 1385 1390 1395

CAA GAA GCA TGT CAT GGT AAT ACT TCA AAT AAA GAA CAG TTA ACT GCT 4461
 Gln Glu Ala Cys His Gly Asn Thr Ser Asn Lys Glu Gln Leu Thr Ala
 1400 1405 1410

ACT AAA ACG GAG CAA AAT ATA AAA GAT TTT GAG ACT TCT GAT ACA TTT 4509
 Thr Lys Thr Glu Gln Asn Ile Lys Asp Phe Glu Thr Ser Asp Thr Phe
 1415 1420 1425

TTT CAG ACT GCA AGT GGG AAA AAT ATT AGT GTC GCC AAA GAG TCA TTT 4557
 Phe Gln Thr Ala Ser Gly Lys Asn Ile Ser Val Ala Lys Glu Ser Phe
 1430 1435 1440

AAT AAA ATT GTA AAT TTC TTT GAT CAG AAA CCA GAA GAA TTG CAT AAC 4605
 Asn Lys Ile Val Asn Phe Phe Asp Gln Lys Pro Glu Glu Leu His Asn
 1445 1450 1455

| | |
|---|------|
| TTT TCC TTA AAT TCT GAA TTA CAT TCT GAC ATA AGA AAG AAC AAA ATG Phe Ser Leu Asn Ser Glu Leu His Ser Asp Ile Arg Lys Asn Lys Met 1460 1465 1470 1475 | 4653 |
| GAC ATT CTA AGT TAT GAG GAA ACA GAC ATA GTT AAA CAC AAA ATA CTG Asp Ile Leu Ser Tyr Glu Glu Thr Asp Ile Val Lys His Lys Ile Leu 1480 1485 1490 | 4701 |
| AAA GAA AGT GTC CCA GTT GGT ACT GGA AAT CAA CTA GTG ACC TTC CAG Lys Glu Ser Val Pro Val Gly Thr Gly Asn Gln Leu Val Thr Phe Gln 1495 1500 1505 | 4749 |
| GGA CAA CCC GAA CGT GAT GAA AAG ATC AAA GAA CCT ACT CTG TTG GGT Gly Gln Pro Glu Arg Asp Glu Lys Ile Lys Glu Pro Thr Leu Leu Gly 1510 1515 1520 | 4797 |
| TTT CAT ACA GCT AGC GGG AAA AAA GTT AAA ATT GCA AAG GAA TCT TTG Phe His Thr Ala Ser Gly Lys Val Lys Ile Ala Lys Glu Ser Leu 1525 1530 1535 | 4845 |
| GAC AAA GTG AAA AAC CTT TTT GAT GAA AAA GAG CAA GGT ACT AGT GAA Asp Lys Val Lys Asn Leu Phe Asp Glu Lys Glu Gln Gly Thr Ser Glu 1540 1545 1550 1555 | 4893 |
| ATC ACC AGT TTT AGC CAT CAA TGG GCA AAG ACC CTA AAG TAC AGA GAG Ile Thr Ser Phe Ser His Gln Trp Ala Lys Thr Leu Lys Tyr Arg Glu 1560 1565 1570 | 4941 |
| GCC TGT AAA GAC CTT GAA TTA GCA TGT GAG ACC ATT GAG ATC ACA GCT Ala Cys Lys Asp Leu Glu Leu Ala Cys Glu Thr Ile Glu Ile Thr Ala 1575 1580 1585 | 4989 |
| GCC CCA AAG TGT AAA GAA ATG CAG AAT TCT CTC AAT AAT GAT AAA AAC Ala Pro Lys Cys Lys Glu Met Gln Asn Ser Leu Asn Asp Lys Asn 1590 1595 1600 | 5037 |
| CTT GTT TCT ATT GAG ACT GTG GTG CCA CCT AAG CTC ATT GAT AAT Leu Val Ser Ile Glu Thr Val Val Pro Pro Lys Leu Leu Ser Asp Asn 1605 1610 1615 | 5085 |
| TTA TGT AGA CAA ACT GAA AAT CTC AAA ACA TCA AAA AGT ATC TTT TTG Leu Cys Arg Gln Thr Glu Asn Leu Lys Thr Ser Lys Ser Ile Phe Leu 1620 1625 1630 1635 | 5133 |
| AAA GTT AAA GTA CAT GAA AAT GTA GAA AAA GAA ACA GCA AAA AGT CCT Lys Val Lys Val His Glu Asn Val Glu Lys Glu Thr Ala Lys Ser Pro 1640 1645 1650 | 5181 |
| GCA ACT TGT TAC ACA AAT CAG TCC CCT TAT TCA GTC ATT GAA AAT TCA Ala Thr Cys Tyr Thr Asn Gln Ser Pro Tyr Ser Val Ile Glu Asn Ser 1655 1660 1665 | 5229 |
| GCC TTA GCT TTT TAC ACA AGT TGT AGT AGA AAA ACT TCT GTG AGT CAG Ala Leu Ala Phe Tyr Thr Ser Cys Ser Arg Lys Thr Ser Val Ser Gln 1670 1675 1680 | 5277 |
| ACT TCA TTA CTT GAA GCA AAA AAA TGG CTT AGA GAA GGA ATA TTT GAT Thr Ser Leu Leu Glu Ala Lys Lys Trp Leu Arg Glu Gly Ile Phe Asp 1685 1690 1695 | 5325 |
| GGT CAA CCA GAA AGA ATA AAT ACT GCA GAT TAT GTA GGA AAT TAT TTG | 5373 |

Gly Gln Pro Glu Arg Ile Asn Thr Ala Asp Tyr Val Gly Asn Tyr Leu
 1700 1705 1710 1715

TAT GAA AAT AAT TCA AAC AGT ACT ATA GCT GAA AAT GAC AAA AAT CAT 5421
 Tyr Glu Asn Asn Ser Asn Ser Thr Ile Ala Glu Asn Asp Lys Asn His
 1720 1725 1730

CTC TCC GAA AAA CAA GAT ACT TAT TTA AGT AAC AGT AGC ATG TCT AAC 5469
 Leu Ser Glu Lys Gln Asp Thr Tyr Leu Ser Asn Ser Met Ser Asn
 1735 1740 1745

AGC TAT TCC TAC CAT TCT GAT GAG GTA TAT AAT GAT TCA GGA TAT CTC 5517
 Ser Tyr Ser Tyr His Ser Asp Glu Val Tyr Asn Asp Ser Gly Tyr Leu
 1750 1755 1760

TCA AAA AAT AAA CTT GAT TCT GGT ATT GAG CCA GTA TTG AAG AAT GTT 5565
 Ser Lys Asn Lys Leu Asp Ser Gly Ile Glu Pro Val Leu Lys Asn Val
 1765 1770 1775

GAA GAT CAA AAA AAC ACT AGT TTT TCC AAA GTA ATA TCC AAT GTA AAA 5613
 Glu Asp Gln Lys Asn Thr Ser Phe Ser Lys Val Ile Ser Asn Val Lys
 1780 1785 1790 1795

GAT GCA AAT GCA TAC CCA CAA ACT GTA AAT GAA GAT ATT TGC GTT GAG 5661
 Asp Ala Asn Ala Tyr Pro Gln Thr Val Asn Glu Asp Ile Cys Val Glu
 1800 1805 1810

GAA CTT GTG ACT AGC TCT TCA CCC TGC AAA AAT AAA AAT GCA GCC ATT 5709
 Glu Leu Val Thr Ser Ser Pro Cys Lys Asn Lys Asn Ala Ala Ile
 1815 1820 1825

AAA TTG TCC ATA TCT AAT AGT AAT AAT TTT GAG GTA GGG CCA CCT GCA 5757
 Lys Leu Ser Ile Ser Asn Ser Asn Phe Glu Val Gly Pro Pro Ala
 1830 1835 1840

TTT AGG ATA GCC AGT GGT AAA ATC GTT TGT GTT TCA CAT GAA ACA ATT 5805
 Phe Arg Ile Ala Ser Gly Ile Val Cys Val Ser His Glu Thr Ile
 1845 1850 1855

AAA AAA GTG AAA GAC ATA TTT ACA GAC AGT TTC AGT AAA GTA ATT AAG 5853
 Lys Lys Val Lys Asp Ile Phe Thr Asp Ser Phe Ser Lys Val Ile Lys
 1860 1865 1870 1875

GAA AAC AAC GAG AAT AAA TCA AAA ATT TGC CAA ACG AAA ATT ATG GCA 5901
 Glu Asn Asn Glu Asn Lys Ser Lys Ile Cys Gln Thr Lys Ile Met Ala
 1880 1885 1890

GGT TGT TAC GAG GCA TTG GAT GAT TCA GAG GAT ATT CTT CAT AAC TCT 5949
 Gly Cys Tyr Glu Ala Leu Asp Asp Ser Glu Asp Ile Leu His Asn Ser
 1895 1900 1905

CTA GAT AAT GAT GAA TGT AGC ACG CAT TCA CAT AAG GTT TTT GCT GAC 5997
 Leu Asp Asn Asp Glu Cys Ser Thr His Ser His Lys Val Phe Ala Asp
 1910 1915 1920

ATT CAG AGT GAA GAA ATT TTA CAA CAT AAC CAA AAT ATG TCT GGA TTG 6045
 Ile Gln Ser Glu Glu Ile Leu Gln His Asn Gln Asn Met Ser Gly Leu
 1925 1930 1935

GAG AAA GTT TCT AAA ATA TCA CCT TGT GAT GTT AGT TTG GAA ACT TCA 6093
 Glu Lys Val Ser Lys Ile Ser Pro Cys Asp Val Ser Leu Glu Thr Ser

| 1940 | 1945 | 1950 | 1955 | |
|--|------|------|------|------|
| GAT ATA TGT AAA TGT AGT ATA GGG AAG CTT CAT AAG TCA GTC TCA TCT Asp Ile Cys Lys Cys Ser Ile Gly Lys Leu His Lys Ser Val Ser Ser | | | | 6141 |
| 1960 | 1965 | 1970 | | |
| GCA AAT ACT TGT GGG ATT TTT AGC ACA GCA AGT GGA AAA TCT GTC CAG Ala Asn Thr Cys Gly Ile Phe Ser Thr Ala Ser Gly Lys Ser Val Gln | 1975 | 1980 | 1985 | 6189 |
| Val Ser Asp Ala Ser Leu Gln Asn Ala Arg Gln Val Phe Ser Glu Ile | | | | |
| 1990 | 1995 | 2000 | | 6237 |
| GAA GAT AGT ACC AAG CAA GTC TTT TCC AAA GTA TTG TTT AAA AGT AAC Glu Asp Ser Thr Lys Gln Val Phe Ser Lys Val Leu Phe Lys Ser Asn | 2005 | 2010 | 2015 | 6285 |
| Glu His Ser Asp Gln Leu Thr Arg Glu Asn Thr Ala Ile Arg Thr | 2020 | 2025 | 2030 | 2035 |
| 2035 | | | | 6333 |
| CCA GAA CAT TTA ATA TCC CAA AAA GGC TTT TCA TAT AAT GTG GTA AAT Pro Glu His Leu Ile Ser Gln Lys Gly Phe Ser Tyr Asn Val Val Asn | 2040 | 2045 | 2050 | 6381 |
| 2050 | | | | |
| TCA TCT GCT TTC TCT GGA TTT AGT ACA GCA AGT GGA AAG CAA GTT TCC Ser Ser Ala Phe Ser Gly Phe Ser Thr Ala Ser Gly Lys Gln Val Ser | 2055 | 2060 | 2065 | 6429 |
| Ile Leu Glu Ser Ser Leu His Lys Val Lys Gly Val Leu Glu Glu Phe | 2070 | 2075 | 2080 | 6477 |
| 2080 | | | | |
| GAT TTA ATC AGA ACT GAG CAT AGT CTT CAC TAT TCA CCT ACG TCT AGA Asp Leu Ile Arg Thr Glu His Ser Leu His Tyr Ser Pro Thr Ser Arg | 2085 | 2090 | 2095 | 6525 |
| 2095 | | | | |
| CAA AAT GTA TCA AAA ATA CTT CCT CGT GTT GAT AAG AGA AAC CCA GAG Gln Asn Val Ser Lys Ile Leu Pro Arg Val Asp Lys Arg Asn Pro Glu | 2100 | 2105 | 2110 | 6573 |
| 2110 | | | | |
| CAC TGT GTA AAC TCA GAA ATG GAA AAA ACC TGC AGT AAA GAA TTT AAA His Cys Val Asn Ser Glu Met Glu Lys Thr Cys Ser Lys Glu Phe Lys | 2120 | 2125 | 2130 | 6621 |
| 2130 | | | | |
| TTA TCA AAT AAC TTA AAT GTT GAA GGT GGT TCT TCA GAA AAT AAT CAC Leu Ser Asn Asn Leu Asn Val Glu Gly Ser Ser Glu Asn Asn His | 2135 | 2140 | 2145 | 6669 |
| 2145 | | | | |
| TCT ATT AAA GTT TCT CCA TAT CTC TCT CAA TTT CAA CAA GAC AAA CAA Ser Ile Lys Val Ser Pro Tyr Leu Ser Gln Phe Gln Gln Asp Lys Gln | 2150 | 2155 | 2160 | 6717 |
| 2160 | | | | |
| CAG TTG GTA TTA GGA ACC AAA GTC TCA CTT GTT GAG AAC ATT CAT GTT Gln Leu Val Leu Gly Thr Lys Val Ser Leu Val Glu Asn Ile His Val | 2165 | 2170 | 2175 | 6765 |
| 2175 | | | | |
| TTG GGA AAA GAA CAG GCT TCA CCT AAA AAC GTA AAA ATG GAA ATT GGT Leu Gly Lys Glu Gln Ala Ser Pro Lys Asn Val Lys Met Glu Ile GLY | 2180 | 2185 | 2190 | 6813 |
| 2190 | | | | |
| 2195 | | | | |

AAA ACT GAA ACT TTT TCT GAT GTT CCT GTG AAA ACA AAT ATA GAA GTT 6861
 Lys Thr Glu Thr Phe Ser Asp Val Pro Val Lys Thr Asn Ile Glu Val
 2200 2205 2210

 TGT TCT ACT TAC TCC AAA GAT TCA GAA AAC TAC TTT GAA ACA GAA GCA 6909
 Cys Ser Thr Tyr Ser Lys Asp Ser Glu Asn Tyr Phe Glu Thr Glu Ala
 2215 2220 2225

 GTA GAA ATT GCT AAA GCT TTT ATG GAA GAT GAT GAA CTG ACA GAT TCT 6957
 Val Glu Ile Ala Lys Ala Phe Met Glu Asp Asp Glu Leu Thr Asp Ser
 2230 2235 2240

 AAA CTG CCA AGT CAT GCC ACA CAT TCT CTT TTT ACA TGT CCC GAA AAT 7005
 Lys Leu Pro Ser His Ala Thr His Ser Leu Phe Thr Cys Pro Glu Asn
 2245 2250 2255

 GAG GAA ATG GTT TTG TCA AAT TCA AGA ATT GGA AAA AGA AGA GGA GAG 7053
 Glu Glu Met Val Leu Ser Asn Ser Arg Ile Gly Lys Arg Arg Gly Glu
 2260 2265 2270 2275

 CCC CTT ATC TTA GTG GGA GAA CCC TCA ATC AAA AGA AAC TTA TTA AAT 7101
 Pro Leu Ile Leu Val Gly Glu Pro Ser Ile Lys Arg Asn Leu Leu Asn
 2280 2285 2290

D
C
X
 GAA TTT GAC AGG ATA ATA GAA AAT CAA GAA AAA TCC TTA AAG GCT TCA 7149
 Glu Phe Asp Arg Ile Ile Glu Asn Gln Glu Lys Ser Leu Lys Ala Ser
 2295 2300 2305

 AAA AGC ACT CCA GAT GGC ACA ATA AAA GAT CGA AGA TTG TTT ATG CAT 7197
 Lys Ser Thr Pro Asp Gly Thr Ile Lys Asp Arg Arg Leu Phe Met His
 2310 2315 2320

 CAT GTT TCT TTA GAG CCG ATT ACC TGT GTA CCC TTT CGC ACA ACT AAG 7245
 His Val Ser Leu Glu Pro Ile Thr Cys. Val Pro Phe Arg Thr Thr Lys
 2325 2330 2335

 GAA CGT CAA GAG ATA CAG AAT CCA AAT TTT ACC GCA CCT GGT CAA GAA 7293
 Glu Arg Gln Glu Ile Gln Asn Pro Asn Phe Thr Ala Pro Gly Gln Glu
 2340 2345 2350 2355

 TTT CTG TCT AAA TCT CAT TTG TAT GAA CAT CTG ACT TTG GAA AAA TCT 7341
 Phe Leu Ser Lys Ser His Leu Tyr Glu His Leu Thr Leu Glu Lys Ser
 2360 2365 2370

 TCA AGC AAT TTA GCA GTT TCA GGA CAT CCA TTT TAT CAA GTT TCT GCT 7389
 Ser Ser Asn Leu Ala Val Ser Gly His Pro Phe Tyr Gln Val Ser Ala
 2375 2380 2385

 ACA AGA AAT GAA AAA ATG AGA CAC TTG ATT ACT ACA GGC AGA CCA ACC 7437
 Thr Arg Asn Glu Lys Met Arg His Leu Ile Thr Thr Gly Arg Pro Thr
 2390 2395 2400

 AAA GTC TTT GTT CCA CCT TTT AAA ACT AAA TCG CAT TTT CAC AGA GTT 7485
 Lys Val Phe Val Pro Pro Phe Lys Thr Lys Ser His Phe His Arg Val
 2405 2410 2415

 GAA CAG TGT GTT AGG AAT ATT AAC TTG GAG GAA AAC AGA CAA AAG CAA 7533
 Glu Gln Cys Val Arg Asn Ile Asn Leu Glu Asn Arg Gln Lys Gln
 2420 2425 2430 2435

D
D
C
M

| | |
|---|------|
| AAC ATT GAT GGA CAT GGC TCT GAT GAT AGT AAA AAT AAG ATT AAT GAC Asn Ile Asp Gly His Gly Ser Asp Asp Ser Lys Asn Lys Ile Asn Asp 2440 2445 2450 | 7581 |
| AAT GAG ATT CAT CAG TTT AAC AAA AAC AAC TCC AAT CAA GCA GCA GCT Asn Glu Ile His Gln Phe Asn Lys Asn Ser Asn Gln Ala Ala Ala 2455 2460 2465 | 7629 |
| GTA ACT TTC ACA AAG TGT GAA GAA CCT TTA GAT TTA ATT ACA AGT Val Thr Phe Thr Lys Cys Glu Glu Pro Leu Asp Leu Ile Thr Ser 2470 2475 2480 | 7677 |
| CTT CAG AAT GCC AGA GAT ATA CAG GAT ATG CGA ATT AAG AAG AAA CAA Leu Gln Asn Ala Arg Asp Ile Gln Asp Met Arg Ile Lys Lys Lys Gln 2485 2490 2495 | 7725 |
| AGG CAA CGC GTC TTT CCA CAG CCA GGC AGT CTG TAT CTT GCA AAA ACA Arg Gln Arg Val Phe Pro Gln Pro Gly Ser Leu Tyr Leu Ala Lys Thr 2500 2505 2510 2515 | 7773 |
| TCC ACT CTG CCT CGA ATC TCT CTG AAA GCA GCA GTA GGA GGC CAA GTT Ser Thr Leu Pro Arg Ile Ser Leu Lys Ala Ala Val Gly Gly Gln Val 2520 2525 2530 | 7821 |
| CCC TCT GCG TGT TCT CAT AAA CAG CTG TAT ACG TAT GGC GTT TCT AAA Pro Ser Ala Cys Ser His Lys Gln Leu Tyr Thr Tyr Gly Val Ser Lys 2535 2540 2545 | 7869 |
| CAT TGC ATA AAA ATT AAC AGC AAA AAT GCA GAG TCT TTT CAG TTT CAC His Cys Ile Lys Ile Asn Ser Lys Asn Ala Glu Ser Phe Gln Phe His 2550 2555 2560 | 7917 |
| ACT GAA GAT TAT TTT GGT AAG GAA AGT TTA TGG ACT GGA AAA GGA ATA Thr Glu Asp Tyr Phe Gly Lys Glu Ser Leu Trp Thr Gly Lys Gly Ile 2565 2570 2575 | 7965 |
| CAG TTG GCT GAT GGT GGA TGG CTC ATA CCC TCC AAT GAT GGA AAG GCT Gln Leu Ala Asp Gly Gly Trp Leu Ile Pro Ser Asn Asp Gly Lys Ala 2580 2585 2590 2595 | 8013 |
| GGA AAA GAA GAA TTT TAT AGG GCT CTG TGT GAC ACT CCA GGT GTG GAT Gly Lys Glu Glu Phe Tyr Arg Ala Leu Cys Asp Thr Pro Gly Val Asp 2600 2605 2610 | 8061 |
| CCA AAG CTT ATT TCT AGA ATT TGG GTT TAT AAT CAC TAT AGA TGG ATC Pro Lys Leu Ile Ser Arg Ile Trp Val Tyr Asn His Tyr Arg Trp Ile 2615 2620 2625 | 8109 |
| ATA TGG AAA CTG GCA GCT ATG GAA TGT GCC TTT CCT AAG GAA TTT GCT Ile Trp Lys Leu Ala Ala Met Glu Cys Ala Phe Pro Lys Glu Phe Ala 2630 2635 2640 | 8157 |
| AAT AGA TGC CTA AGC CCA GAA AGG GTG CTT CTT CAA CTA AAA TAC AGA Asn Arg Cys Leu Ser Pro Glu Arg Val Leu Leu Gln Leu Lys Tyr Arg 2645 2650 2655 | 8205 |
| TAT GAT ACG GAA ATT GAT AGA AGC AGA AGA TCG GCT ATA AAA AAG ATA Tyr Asp Thr Glu Ile Asp Arg Ser Arg Arg Ser Ala Ile Lys Lys Ile 2660 2665 2670 2675 | 8253 |
| ATG GAA AGG GAT GAC ACA GCT GCA AAA ACA CTT GTT CTC TGT GTT TCT | 8301 |

Met Glu Arg Asp Asp Thr Ala Ala Lys Thr Leu Val Leu Cys Val Ser
 2680 2685 2690

GAC ATA ATT TCA TTG AGC GCA AAT ATA TCT GAA ACT TCT AGC AAT AAA 8349
 Asp Ile Ile Ser Leu Ser Ala Asn Ile Ser Glu Thr Ser Ser Asn Lys
 2695 2700 2705

ACT AGT AGT GCA GAT ACC CAA AAA GTG GCC ATT ATT GAA CTT ACA GAT 8397
 Thr Ser Ser Ala Asp Thr Gln Lys Val Ala Ile Ile Glu Leu Thr Asp
 2710 2715 2720

GGG TGG TAT GCT GTT AAG GCC CAG TTA GAT CCT CCC CTC TTA GCT GTC 8445
 Gly Trp Tyr Ala Val Lys Ala Gln Leu Asp Pro Pro Leu Leu Ala Val
 2725 2730 2735

TTA AAG AAT GGC AGA CTG ACA GTT GGT CAG AAG ATT ATT CTT CAT GGA 8493
 Leu Lys Asn Gly Arg Leu Thr Val Gly Gln Lys Ile Ile Leu His Gly
 2740 2745 2750 2755

GCA GAA CTG GTG GGC TCT CCT GAT GCC TGT ACA CCT CTT GAA GCC CCA 8541
 Ala Glu Leu Val Gly Ser Pro Asp Ala Cys Thr Pro Leu Glu Ala Pro
 2760 2765 2770

GAA TCT CTT ATG TTA AAG ATT TCT GCT AAC AGT ACT CGG CCT GCT CGC 8589
 Glu Ser Leu Met Leu Lys Ile Ser Ala Asn Ser Thr Arg Pro Ala Arg
 2775 2780 2785

TGG TAT ACC AAA CTT GGA TTC TTT CCT GAC CCT AGA CCT TTT CCT CTG 8637
 Trp Tyr Thr Lys Leu Gly Phe Phe Pro Asp Pro Arg Pro Phe Pro Leu
 2790 2795 2800

CCC TTA TCA TCG CTT TTC AGT GAT GGA GGA AAT GTT GGT TGT GTT GAT 8685
 Pro Leu Ser Ser Leu Phe Ser Asp Gly Gly Asn Val Gly Cys Val Asp
 2805 2810 2815

GTA ATT ATT CAA AGA GCA TAC CCT ATA CAG TGG ATG GAG AAG ACA TCA 8733
 Val Ile Ile Gln Arg Ala Tyr Pro Ile Gln Trp Met Glu Lys Thr Ser
 2820 2825 2830 2835

TCT GGA TTA TAC ATA TTT CGC AAT GAA AGA GAG GAA GAA AAG GAA GCA 8781
 Ser Gly Leu Tyr Ile Phe Arg Asn Glu Arg Glu Glu Glu Lys Glu Ala
 2840 2845 2850

GCA AAA TAT GTG GAG GCC CAA CAA AAG AGA CTA GAA GCC TTA TTC ACT 8829
 Ala Lys Tyr Val Glu Ala Gln Gln Lys Arg Leu Glu Ala Leu Phe Thr
 2855 2860 2865

AAA ATT CAG GAG GAA TTT GAA GAA CAT GAA GAA AAC ACA ACA AAA CCA 8877
 Lys Ile Gln Glu Glu Phe Glu Glu His Glu Glu Asn Thr Thr Lys Pro
 2870 2875 2880

TAT TTA CCA TCA CGT GCA CTA ACA AGA CAG CAA GTT CGT GCT TTG CAA 8925
 Tyr Leu Pro Ser Arg Ala Leu Thr Arg Gln Gln Val Arg Ala Leu Gln
 2885 2890 2895

GAT GGT GCA GAG CTT TAT GAA GCA GTG AAG AAT GCA GCA GAC CCA GCT 8973
 Asp Gly Ala Glu Leu Tyr Glu Ala Val Lys Asn Ala Ala Asp Pro Ala
 2900 2905 2910 2915

TAC CTT GAG GGT TAT TTC AGT GAA GAG CAG TTA AGA GCC TTG AAT AAT 9021
 Tyr Leu Glu Gly Tyr Phe Ser Glu Glu Gln Leu Arg Ala Leu Asn Asn

| 2920 | 2925 | 2930 | |
|--|------|------|------|
| CAC AGG CAA ATG TTG AAT GAT AAG AAA CAA GCT CAG ATC CAG TTG GAA His Arg Gln Met Leu Asn Asp Lys Lys Gln Ala Gln Ile Gln Leu Glu 2935 | 2940 | 2945 | 9069 |
| ATT AGG AAG GCC ATG GAA TCT GCT GAA CAA AAG GAA CAA GGT TTA TCA Ile Arg Lys Ala Met Glu Ser Ala Glu Gln Lys Glu Gln Gly Leu Ser 2950 | 2955 | 2960 | 9117 |
| AGG GAT GTC ACA ACC GTG TGG AAG TTG CGT ATT GTA AGC TAT TCA AAA Arg Asp Val Thr Thr Val Trp Lys Leu Arg Ile Val Ser Tyr Ser Lys 2965 | 2970 | 2975 | 9165 |
| AAA GAA AAA GAT TCA GTT ATA CTG AGT ATT TGG CGT CCA TCA TCA GAT Lys Glu Lys Asp Ser Val Ile Leu Ser Ile Trp Arg Pro Ser Ser Asp 2980 | 2985 | 2990 | 9213 |
| TTA TAT TCT CTG TTA ACA GAA GGA AAG AGA TAC AGA ATT TAT CAT CTT Leu Tyr Ser Leu Leu Thr Glu Gly Lys Arg Tyr Arg Ile Tyr His Leu 3000 | 3005 | 3010 | 9261 |
| GCA ACT TCA AAA TCT AAA AGT AAA TCT GAA AGA GCT AAC ATA CAG TTA Ala Thr Ser Lys Ser Lys Ser Glu Arg Ala Asn Ile Gln Leu 3015 | 3020 | 3025 | 9309 |
| GCA GCG ACA AAA AAA ACT CAG TAT CAA CAA CTA CCG GTT TCA GAT GAA Ala Ala Thr Lys Lys Thr Gln Tyr Gln Gln Leu Pro Val Ser Asp Glu 3030 | 3035 | 3040 | 9357 |
| ATT TTA TTT CAG ATT TAC CAG CCA CGG GAG CCC CTT CAC TTC AGC AAA Ile Leu Phe Gln Ile Tyr Gln Pro Arg Glu Pro Leu His Phe Ser Lys 3045 | 3050 | 3055 | 9405 |
| TTT TTA GAT CCA GAC TTT CAG CCA TCT TGT TCT GAG GTG GAC CTA ATA Phe Leu Asp Pro Asp Phe Gln Pro Ser Cys Ser Glu Val Asp Leu Ile 3060 | 3065 | 3070 | 9453 |
| GGA TTT GTC GTT TCT GTT GTG AAA AAA ACA GGA CTT GCC CCT TTC GTC Gly Phe Val Val Ser Val Val Lys Lys Thr Gly Leu Ala Pro Phe Val 3080 | 3085 | 3090 | 9501 |
| TAT TTG TCA GAC GAA TGT TAC AAT TTA CTG GCA ATA AAG TTT TGG ATA Tyr Leu Ser Asp Glu Cys Tyr Asn Leu Leu Ala Ile Lys Phe Trp Ile 3095 | 3100 | 3105 | 9549 |
| GAC CTT AAT GAG GAC ATT ATT AAG CCT CAT ATG TTA ATT GCT GCA AGC Asp Leu Asn Glu Asp Ile Ile Lys Pro His Met Leu Ile Ala Ala Ser 3110 | 3115 | 3120 | 9597 |
| AAC CTC CAG TGG CGA CCA GAA TCC AAA TCA GGC CTT CTT ACT TTA TTT Asn Leu Gln Trp Arg Pro Glu Ser Lys Ser Gly Leu Leu Thr Leu Phe 3125 | 3130 | 3135 | 9645 |
| GCT GGA GAT TTT TCT GTG TTT TCT GCT AGT CCA AAA GAG GGC CAC TTT Ala Gly Asp Phe Ser Val Phe Ser Ala Ser Pro Lys Glu Gly His Phe 3140 | 3145 | 3150 | 9693 |
| CAA GAG ACA TTC AAC AAA ATG AAA AAT ACT GTT GAG AAT ATT GAC ATA Gln Glu Thr Phe Asn Lys Met Lys Asn Thr Val Glu Asn Ile Asp Ile 3160 | 3165 | 3170 | 9741 |

CTT TGC AAT GAA GCA GAA AAC AAG CTT ATG CAT ATA CTG CAT GCA AAT 9789
 Leu Cys Asn Glu Ala Glu Asn Lys Leu Met His Ile Leu His Ala Asn
 3175 3180 3185

GAT CCC AAG TGG TCC ACC CCA ACT AAA GAC TGT ACT TCA GGG CCG TAC 9837
 Asp Pro Lys Trp Ser Thr Pro Thr Lys Asp Cys Thr Ser Gly Pro Tyr
 3190 3195 3200

ACT GCT CAA ATC ATT CCT GGT ACA GGA AAC AAG CTT CTG ATG TCT TCT 9885
 Thr Ala Gln Ile Ile Pro Gly Thr Gly Asn Lys Leu Leu Met Ser Ser
 3205 3210 3215

CCT AAT TGT GAG ATA TAT TAT CAA AGT CCT TTA TCA CTT TGT ATG GCC 9933
 Pro Asn Cys Glu Ile Tyr Tyr Gln Ser Pro Leu Ser Leu Cys Met Ala
 3220 3225 3230 3235

AAA AGG AAG TCT GTT TCC ACA CCT GTC TCA GCC CAG ATG ACT TCA AAG 9981
 Lys Arg Lys Ser Val Ser Thr Pro Val Ser Ala Gln Met Thr Ser Lys
 3240 3245 3250

TCT TGT AAA GGG GAG AAA GAG ATT GAT GAC CAA AAG AAC TGC AAA AAG 10029
 Ser Cys Lys Gly Glu Lys Glu Ile Asp Asp Gln Lys Asn Cys Lys Lys
 3255 3260 3265

D X
 AGA AGA GCC TTG GAT TTC TTG AGT AGA CTG CCT TTA CCT CCA CCT GTT 10077
 Arg Arg Ala Leu Asp Phe Leu Ser Arg Leu Pro Leu Pro Pro Val
 3270 3275 3280

CCA CCA AGG AGT TGT GGC ACC AAA TAC GAA ACA CCC ATA AAG AAA AAA 10125
 Pro Pro Arg Ser Cys Gly Thr Lys Tyr Glu Thr Pro Ile Lys Lys Lys
 3300 3305 3310 3315

GAA CTG AAT TCT CCT CAG ATG ACT CCA TTT AAA AAA TTC AAT GAA ATT 10221
 Glu Leu Asn Ser Pro Gln Met Thr Pro Phe Lys Lys Phe Asn Glu Ile
 3320 3325 3330

TCT CTT TTG GAA AGT AAT TCA ATA GCT GAC GAA GAA CTT GCA TTG ATA 10269
 Ser Leu Leu Glu Ser Asn Ser Ile Ala Asp Glu Glu Leu Ala Leu Ile
 3335 3340 3345

AAT ACC CAA GCT CTT TTG TCT GGT TCA ACA GGA GAA AAA CAA TTT ATA 10317
 Asn Thr Gln Ala Leu Leu Ser Gly Ser Thr Gly Glu Lys Gln Phe Ile
 3350 3355 3360

TCT GTC AGT GAA TCC ACT AGG ACT GCT CCC ACC AGT TCA GAA GAT TAT 10365
 Ser Val Ser Glu Ser Thr Arg Thr Ala Pro Thr Ser Glu Asp Tyr
 3365 3370 3375

CTC AGA CTG AAA CGA CGT TGT ACT ACA TCT CTG ATC AAA GAA CAG GAG 10413
 Leu Arg Leu Lys Arg Arg Cys Thr Thr Ser Leu Ile Lys Glu Gln Glu
 3380 3385 3390 3395

AGT TCC CAG GCC AGT ACG GAA GAA TGT GAG AAA AAT AAG CAG GAC ACA 10461
 Ser Ser Gln Ala Ser Thr Glu Glu Cys Glu Lys Asn Lys Gln Asp Thr
 3400 3405 3410

ATT ACA ACT AAA AAA TAT ATC TAA
Ile Thr Thr Lys Lys Tyr Ile
3415

10485

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 3418 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Met Pro Ile Gly Ser Lys Glu Arg Pro Thr Phe Phe Glu Ile Phe Lys
1 5 10 15
Thr Arg Cys Asn Lys Ala Asp Leu Gly Pro Ile Ser Leu Asn Trp Phe
20 25 30
Glu Glu Leu Ser Ser Glu Ala Pro Pro Tyr Asn Ser Glu Pro Ala Glu
35 40 45
Glu Ser Glu His Lys Asn Asn Tyr Glu Pro Asn Leu Phe Lys Thr
50 55 60
Pro Gln Arg Lys Pro Ser Tyr Asn Gln Leu Ala Ser Thr Pro Ile Ile
65 70 75 80
Phe Lys Glu Gln Gly Leu Thr Leu Pro Leu Tyr Gln Ser Pro Val Lys
85 90 95
Glu Leu Asp Lys Phe Lys Leu Asp Leu Gly Arg Asn Val Pro Asn Ser
100 105 110
Arg His Lys Ser Leu Arg Thr Val Lys Thr Lys Met Asp Gln Ala Asp
115 120 125
Asp Val Ser Cys Pro Leu Leu Asn Ser Cys Leu Ser Glu Ser Pro Val
130 135 140
Val Leu Gln Cys Thr His Val Thr Pro Gln Arg Asp Lys Ser Val Val
145 150 155 160
Cys Gly Ser Leu Phe His Thr Pro Lys Phe Val Lys Gly Arg Gln Thr
165 170 175
Pro Lys His Ile Ser Glu Ser Leu Gly Ala Glu Val Asp Pro Asp Met
180 185 190
Ser Trp Ser Ser Ser Leu Ala Thr Pro Pro Thr Leu Ser Ser Thr Val
195 200 205
Leu Ile Val Arg Asn Glu Glu Ala Ser Glu Thr Val Phe Pro His Asp
210 215 220
Thr Thr Ala Asn Val Lys Ser Tyr Phe Ser Asn His Asp Glu Ser Leu
225 230 235 240
Lys Lys Asn Asp Arg Phe Ile Ala Ser Val Thr Asp Ser Glu Asn Thr
245 250 255
Asn Gln Arg Glu Ala Ala Ser His Gly Phe Gly Lys Thr Ser Gly Asn
260 265 270
Ser Phe Lys Val Asn Ser Cys Lys Asp His Ile Gly Lys Ser Met Pro
275 280 285
His Val Leu Glu Asp Glu Val Tyr Glu Thr Val Val Asp Thr Ser Glu
290 295 300
Glu Asp Ser Phe Ser Leu Cys Phe Ser Lys Cys Arg Thr Lys Asn Leu
305 310 315 320
Gln Lys Val Arg Thr Ser Lys Thr Arg Lys Lys Ile Phe His Glu Ala
325 330 335
Asn Ala Asp Glu Cys Glu Lys Ser Lys Asn Gln Val Lys Glu Lys Tyr

340 345 350
 Ser Phe Val Ser Glu Val Glu Pro Asn Asp Thr Asp Pro Leu Asp Ser
 355 360 365
 Asn Val Ala Asn Gln Lys Pro Phe Glu Ser Gly Ser Asp Lys Ile Ser
 370 375 380
 Lys Glu Val Val Pro Ser Leu Ala Cys Glu Trp Ser Gln Leu Thr Leu
 385 390 395 400
 Ser Gly Leu Asn Gly Ala Gln Met Glu Lys Ile Pro Leu Leu His Ile
 405 410 415
 Ser Ser Cys Asp Gln Asn Ile Ser Glu Lys Asp Leu Leu Asp Thr Glu
 420 425 430
 Asn Lys Arg Lys Lys Asp Phe Leu Thr Ser Glu Asn Ser Leu Pro Arg
 435 440 445
 Ile Ser Ser Leu Pro Lys Ser Glu Lys Pro Leu Asn Glu Glu Thr Val
 450 455 460
 Val Asn Lys Arg Asp Glu Glu Gln His Leu Glu Ser His Thr Asp Cys
 465 470 475 480
 Ile Leu Ala Val Lys Gln Ala Ile Ser Gly Thr Ser Pro Val Ala Ser
 485 490 495
 Ser Phe Gln Gly Ile Lys Lys Ser Ile Phe Arg Ile Arg Glu Ser Pro
 500 505 510
 Lys Glu Thr Phe Asn Ala Ser Phe Ser Gly His Met Thr Asp Pro Asn
 515 520 525
 Phe Lys Lys Glu Thr Glu Ala Ser Glu Ser Gly Leu Glu Ile His Thr
 530 535 540
 Val Cys Ser Gln Lys Glu Asp Ser Leu Cys Pro Asn Leu Ile Asp Asn
 545 550 555 560
 Gly Ser Trp Pro Ala Thr Thr Gln Asn Ser Val Ala Leu Lys Asn
 565 570 575
 Ala Gly Leu Ile Ser Thr Leu Lys Lys Thr Asn Lys Phe Ile Tyr
 580 585 590
 Ala Ile His Asp Glu Thr Ser Tyr Lys Gly Lys Lys Ile Pro Lys Asp
 595 600 605
 Gln Lys Ser Glu Leu Ile Asn Cys Ser Ala Gln Phe Glu Ala Asn Ala
 610 615 620
 Phe Glu Ala Pro Leu Thr Phe Ala Asn Ala Asp Ser Gly Leu Leu His
 625 630 635 640
 Ser Ser Val Lys Arg Ser Cys Ser Gln Asn Asp Ser Glu Glu Pro Thr
 645 650 655
 Leu Ser Leu Thr Ser Ser Phe Gly Thr Ile Leu Arg Lys Cys Ser Arg
 660 665 670
 Asn Glu Thr Cys Ser Asn Asn Thr Val Ile Ser Gln Asp Leu Asp Tyr
 675 680 685
 Lys Glu Ala Lys Cys Asn Lys Glu Lys Leu Gln Leu Phe Ile Thr Pro
 690 695 700
 Glu Ala Asp Ser Leu Ser Cys Leu Gln Glu Gly Gln Cys Glu Asn Asp
 705 710 715 720
 Pro Lys Ser Lys Lys Val Ser Asp Ile Lys Glu Glu Val Leu Ala Ala
 725 730 735
 Ala Cys His Pro Val Gln His Ser Lys Val Glu Tyr Ser Asp Thr Asp
 740 745 750
 Phe Gln Ser Gln Lys Ser Leu Leu Tyr Asp His Glu Asn Ala Ser Thr
 755 760 765
 Leu Ile Leu Thr Pro Thr Ser Lys Asp Val Leu Ser Asn Leu Val Met
 770 775 780
 Ile Ser Arg Gly Lys Glu Ser Tyr Lys Met Ser Asp Lys Leu Lys Gly
 785 790 795 800
 Asn Asn Tyr Glu Ser Asp Val Glu Leu Thr Lys Asn Ile Pro Met Glu
 805 810 815
 Lys Asn Gln Asp Val Cys Ala Leu Asn Glu Asn Tyr Lys Asn Val Glu
 820 825 830

Leu Leu Pro Pro Glu Lys Tyr Met Arg Val Ala Ser Pro Ser Arg Lys
 835 840 845
 Val Gln Phe Asn Gln Asn Thr Asn Leu Arg Val Ile Gln Lys Asn Gln
 850 855 860
 Glu Glu Thr Thr Ser Ile Ser Lys Ile Thr Val Asn Pro Asp Ser Glu
 865 870 875 880
 Glu Leu Phe Ser Asp Asn Glu Asn Asn Phe Val Phe Gln Val Ala Asn
 885 890 895
 Glu Arg Asn Asn Leu Ala Leu Gly Asn Thr Lys Glu Leu His Glu Thr
 900 905 910
 Asp Leu Thr Cys Val Asn Glu Pro Ile Phe Lys Asn Ser Thr Met Val
 915 920 925
 Leu Tyr Gly Asp Thr Gly Asp Lys Gln Ala Thr Gln Val Ser Ile Lys
 930 935 940
 Lys Asp Leu Val Tyr Val Leu Ala Glu Glu Asn Lys Asn Ser Val Lys
 945 950 955 960
 Gln His Ile Lys Met Thr Leu Gly Gln Asp Leu Lys Ser Asp Ile Ser
 965 970 975
 Leu Asn Ile Asp Lys Ile Pro Glu Lys Asn Asn Asp Tyr Met Asp Lys
 980 985 990
 Trp Ala Gly Leu Leu Gly Pro Ile Ser Asn His Ser Phe Gly Gly Ser
 995 1000 1005
 Phe Arg Thr Ala Ser Asn Lys Glu Ile Lys Leu Ser Glu His Asn Ile
 1010 1015 1020
 Lys Lys Ser Lys Met Phe Phe Lys Asp Ile Glu Glu Gln Tyr Pro Thr
 1025 1030 1035 1040
 Ser Leu Ala Cys Val Glu Ile Val Asn Thr Leu Ala Leu Asp Asn Gln
 1045 1050 1055
 Lys Lys Leu Ser Lys Pro Gln Ser Ile Asn Thr Val Ser Ala His Leu
 1060 1065 1070
 Gln Ser Ser Val Val Val Ser Asp Cys Lys Asn Ser His Ile Thr Pro
 1075 1080 1085
 Gln Met Leu Phe Ser Lys Gln Asp Phe Asn Ser Asn His Asn Leu Thr
 1090 1095 1100
 Pro Ser Gln Lys Ala Glu Ile Thr Glu Leu Ser Thr Ile Leu Glu Glu
 1105 1110 1115 1120
 Ser Gly Ser Gln Phe Glu Phe Thr Gln Phe Arg Lys Pro Ser Tyr Ile
 1125 1130 1135
 Leu Gln Lys Ser Thr Phe Glu Val Pro Glu Asn Gln Met Thr Ile Leu
 1140 1145 1150
 Lys Thr Thr Ser Glu Glu Cys Arg Asp Ala Asp Leu His Val Ile Met
 1155 1160 1165
 Asn Ala Pro Ser Ile Gly Gln Val Asp Ser Ser Lys Gln Phe Glu Gly
 1170 1175 1180
 Thr Val Glu Ile Lys Arg Lys Phe Ala Gly Leu Leu Lys Asn Asp Cys
 1185 1190 1195 1200
 Asn Lys Ser Ala Ser Gly Tyr Leu Thr Asp Glu Asn Glu Val Gly Phe
 1205 1210 1215
 Arg Gly Phe Tyr Ser Ala His Gly Thr Lys Leu Asn Val Ser Thr Glu
 1220 1225 1230
 Ala Leu Gln Lys Ala Val Lys Leu Phe Ser Asp Ile Glu Asn Ile Ser
 1235 1240 1245
 Glu Glu Thr Ser Ala Glu Val His Pro Ile Ser Leu Ser Ser Ser Lys
 1250 1255 1260
 Cys His Asp Ser Val Val Ser Met Phe Lys Ile Glu Asn His Asn Asp
 1265 1270 1275 1280
 Lys Thr Val Ser Glu Lys Asn Asn Lys Cys Gln Leu Ile Leu Gln Asn
 1285 1290 1295
 Asn Ile Glu Met Thr Thr Gly Thr Phe Val Glu Glu Ile Thr Glu Asn
 1300 1305 1310
 Tyr Lys Arg Asn Thr Glu Asn Glu Asp Asn Lys Tyr Thr Ala Ala Ser

| | 1315 | 1320 | 1325 |
|------|---|------|------|
| Arg | Asn Ser His Asn Leu Glu Phe Asp Gly Ser Asp Ser Ser Lys Asn | | |
| | 1330 | 1335 | 1340 |
| Asp | Thr Val Cys Ile His Lys Asp Glu Thr Asp Leu Leu Phe Thr Asp | | |
| 1345 | 1350 | 1355 | 1360 |
| Gln | His Asn Ile Cys Leu Lys Leu Ser Gly Gln Phe Met Lys Glu Gly | | |
| | 1365 | 1370 | 1375 |
| Asn | Thr Gln Ile Lys Glu Asp Leu Ser Asp Leu Thr Phe Leu Glu Val | | |
| | 1380 | 1385 | 1390 |
| Ala | Lys Ala Gln Glu Ala Cys His Gly Asn Thr Ser Asn Lys Glu Gln | | |
| | 1395 | 1400 | 1405 |
| Leu | Thr Ala Thr Lys Thr Glu Gln Asn Ile Lys Asp Phe Glu Thr Ser | | |
| | 1410 | 1415 | 1420 |
| Asp | Thr Phe Phe Gln Thr Ala Ser Gly Lys Asn Ile Ser Val Ala Lys | | |
| 1425 | 1430 | 1435 | 1440 |
| Glu | Ser Phe Asn Lys Ile Val Asn Phe Phe Asp Gln Lys Pro Glu Glu | | |
| | 1445 | 1450 | 1455 |
| Leu | His Asn Phe Ser Leu Asn Ser Glu Leu His Ser Asp Ile Arg Lys | | |
| | 1460 | 1465 | 1470 |
| Asn | Lys Met Asp Ile Leu Ser Tyr Glu Glu Thr Asp Ile Val Lys His | | |
| | 1475 | 1480 | 1485 |
| Lys | Ile Leu Lys Glu Ser Val Pro Val Gly Thr Gly Asn Gln Leu Val | | |
| | 1490 | 1495 | 1500 |
| Thr | Phe Gln Gly Gln Pro Glu Arg Asp Glu Lys Ile Lys Glu Pro Thr | | |
| 1505 | 1510 | 1515 | 1520 |
| Leu | Leu Gly Phe His Thr Ala Ser Gly Lys Lys Val Lys Ile Ala Lys | | |
| | 1525 | 1530 | 1535 |
| Glu | Ser Leu Asp Lys Val Lys Asn Leu Phe Asp Glu Lys Glu Gln Gly | | |
| | 1540 | 1545 | 1550 |
| Thr | Ser Glu Ile Thr Ser Phe Ser His Gln Trp Ala Lys Thr Leu Lys | | |
| | 1555 | 1560 | 1565 |
| Tyr | Arg Glu Ala Cys Lys Asp Leu Glu Leu Ala Cys Glu Thr Ile Glu | | |
| | 1570 | 1575 | 1580 |
| Ile | Thr Ala Ala Pro Lys Cys Lys Glu Met Gln Asn Ser Leu Asn Asn | | |
| 1585 | 1590 | 1595 | 1600 |
| Asp | Lys Asn Leu Val Ser Ile Glu Thr Val Val Pro Pro Lys Leu Leu | | |
| | 1605 | 1610 | 1615 |
| Ser | Asp Asn Leu Cys Arg Gln Thr Glu Asn Leu Lys Thr Ser Lys Ser | | |
| | 1620 | 1625 | 1630 |
| Ile | Phe Leu Lys Val Lys Val His Glu Asn Val Glu Lys Glu Thr Ala | | |
| | 1635 | 1640 | 1645 |
| Lys | Ser Pro Ala Thr Cys Tyr Thr Asn Gln Ser Pro Tyr Ser Val Ile | | |
| | 1650 | 1655 | 1660 |
| Glu | Asn Ser Ala Leu Ala Phe Tyr Thr Ser Cys Ser Arg Lys Thr Ser | | |
| 1665 | 1670 | 1675 | 1680 |
| Val | Ser Gln Thr Ser Leu Leu Glu Ala Lys Lys Trp Leu Arg Glu Gly | | |
| | 1685 | 1690 | 1695 |
| Ile | Phe Asp Gly Gln Pro Glu Arg Ile Asn Thr Ala Asp Tyr Val Gly | | |
| | 1700 | 1705 | 1710 |
| Asn | Tyr Leu Tyr Glu Asn Asn Ser Asn Ser Thr Ile Ala Glu Asn Asp | | |
| | 1715 | 1720 | 1725 |
| Lys | Asn His Leu Ser Glu Lys Gln Asp Thr Tyr Leu Ser Asn Ser Ser | | |
| | 1730 | 1735 | 1740 |
| Met | Ser Asn Ser Tyr Ser Tyr His Ser Asp Glu Val Tyr Asn Asp Ser | | |
| 1745 | 1750 | 1755 | 1760 |
| Gly | Tyr Leu Ser Lys Asn Lys Leu Asp Ser Gly Ile Glu Pro Val Leu | | |
| | 1765 | 1770 | 1775 |
| Lys | Asn Val Glu Asp Gln Lys Asn Thr Ser Phe Ser Lys Val Ile Ser | | |
| | 1780 | 1785 | 1790 |
| Asn | Val Lys Asp Ala Asn Ala Tyr Pro Gln Thr Val Asn Glu Asp Ile | | |
| | 1795 | 1800 | 1805 |

Cys Val Glu Glu Leu Val Thr Ser Ser Ser Pro Cys Lys Asn Lys Asn
 1810 1815 1820
 Ala Ala Ile Lys Leu Ser Ile Ser Asn Ser Asn Asn Phe Glu Val Gly
 1825 1830 1835 1840
 Pro Pro Ala Phe Arg Ile Ala Ser Gly Lys Ile Val Cys Val Ser His
 1845 1850 1855
 Glu Thr Ile Lys Lys Val Lys Asp Ile Phe Thr Asp Ser Phe Ser Lys
 1860 1865 1870
 Val Ile Lys Glu Asn Asn Glu Asn Lys Ser Lys Ile Cys Gln Thr Lys
 1875 1880 1885
 Ile Met Ala Gly Cys Tyr Glu Ala Leu Asp Asp Ser Glu Asp Ile Leu
 1890 1895 1900
 His Asn Ser Leu Asp Asn Asp Glu Cys Ser Thr His Ser His Lys Val
 1905 1910 1915 1920
 Phe Ala Asp Ile Gln Ser Glu Glu Ile Leu Gln His Asn Gln Asn Met
 1925 1930 1935
 Ser Gly Leu Glu Lys Val Ser Lys Ile Ser Pro Cys Asp Val Ser Leu
 1940 1945 1950
 Glu Thr Ser Asp Ile Cys Lys Cys Ser Ile Gly Lys Leu His Lys Ser
 1955 1960 1965
 Val Ser Ser Ala Asn Thr Cys Gly Ile Phe Ser Thr Ala Ser Gly Lys
 1970 1975 1980
 Ser Val Gln Val Ser Asp Ala Ser Leu Gln Asn Ala Arg Gln Val Phe
 1985 1990 1995 2000
 Ser Glu Ile Glu Asp Ser Thr Lys Gln Val Phe Ser Lys Val Leu Phe
 2005 2010 2015
 Lys Ser Asn Glu His Ser Asp Gln Leu Thr Arg Glu Glu Asn Thr Ala
 2020 2025 2030
 Ile Arg Thr Pro Glu His Leu Ile Ser Gln Lys Gly Phe Ser Tyr Asn
 2035 2040 2045
 Val Val Asn Ser Ser Ala Phe Ser Gly Phe Ser Thr Ala Ser Gly Lys
 2050 2055 2060
 Gln Val Ser Ile Leu Glu Ser Ser Leu His Lys Val Lys Gly Val Leu
 2065 2070 2075 2080
 Glu Glu Phe Asp Leu Ile Arg Thr Glu His Ser Leu His Tyr Ser Pro
 2085 2090 2095
 Thr Ser Arg Gln Asn Val Ser Lys Ile Leu Pro Arg Val Asp Lys Arg
 2100 2105 2110
 Asn Pro Glu His Cys Val Asn Ser Glu Met Glu Lys Thr Cys Ser Lys
 2115 2120 2125
 Glu Phe Lys Leu Ser Asn Asn Leu Asn Val Glu Gly Ser Ser Glu
 2130 2135 2140
 Asn Asn His Ser Ile Lys Val Ser Pro Tyr Leu Ser Gln Phe Gln Gln
 2145 2150 2155 2160
 Asp Lys Gln Gln Leu Val Leu Gly Thr Lys Val Ser Leu Val Glu Asn
 2165 2170 2175
 Ile His Val Leu Gly Lys Glu Gln Ala Ser Pro Lys Asn Val Lys Met
 2180 2185 2190
 Glu Ile Gly Lys Thr Glu Thr Phe Ser Asp Val Pro Val Lys Thr Asn
 2195 2200 2205
 Ile Glu Val Cys Ser Thr Tyr Ser Lys Asp Ser Glu Asn Tyr Phe Glu
 2210 2215 2220
 Thr Glu Ala Val Glu Ile Ala Lys Ala Phe Met Glu Asp Asp Glu Leu
 2225 2230 2235 2240
 Thr Asp Ser Lys Leu Pro Ser His Ala Thr His Ser Leu Phe Thr Cys
 2245 2250 2255
 Pro Glu Asn Glu Glu Met Val Leu Ser Asn Ser Arg Ile Gly Lys Arg
 2260 2265 2270
 Arg Gly Glu Pro Leu Ile Leu Val Gly Glu Pro Ser Ile Lys Arg Asn
 2275 2280 2285
 Leu Leu Asn Glu Phe Asp Arg Ile Ile Glu Asn Gln Glu Lys Ser Leu

2290 2295 2300
 Lys Ala Ser Lys Ser Thr Pro Asp Gly Thr Ile Lys Asp Arg Arg Leu
 2305 2310 2315 2320
 Phe Met His His Val Ser Leu Glu Pro Ile Thr Cys Val Pro Phe Arg
 2325 2330 2335
 Thr Thr Lys Glu Arg Gln Glu Ile Gln Asn Pro Asn Phe Thr Ala Pro
 2340 2345 2350
 Gly Gln Glu Phe Leu Ser Lys Ser His Leu Tyr Glu His Leu Thr Leu
 2355 2360 2365
 Glu Lys Ser Ser Ser Asn Leu Ala Val Ser Gly His Pro Phe Tyr Gln
 2370 2375 2380
 Val Ser Ala Thr Arg Asn Glu Lys Met Arg His Leu Ile Thr Thr Gly
 2385 2390 2395 2400
 Arg Pro Thr Lys Val Phe Val Pro Pro Phe Lys Thr Lys Ser His Phe
 2405 2410 2415
 His Arg Val Glu Gln Cys Val Arg Asn Ile Asn Leu Glu Glu Asn Arg
 2420 2425 2430
 Gln Lys Gln Asn Ile Asp Gly His Gly Ser Asp Asp Ser Lys Asn Lys
 2435 2440 2445
 Ile Asn Asp Asn Glu Ile His Gln Phe Asn Lys Asn Asn Ser Asn Gln
 2450 2455 2460
 Ala Ala Ala Val Thr Phe Thr Lys Cys Glu Glu Glu Pro Leu Asp Leu
 2465 2470 2475 2480
 Ile Thr Ser Leu Gln Asn Ala Arg Asp Ile Gln Asp Met Arg Ile Lys
 2485 2490 2495
 Lys Lys Gln Arg Gln Arg Val Phe Pro Gln Pro Gly Ser Leu Tyr Leu
 2500 2505 2510
 Ala Lys Thr Ser Thr Leu Pro Arg Ile Ser Leu Lys Ala Ala Val Gly
 2515 2520 2525
 Gly Gln Val Pro Ser Ala Cys Ser His Lys Gln Leu Tyr Thr Tyr Gly
 2530 2535 2540
 Val Ser Lys His Cys Ile Lys Ile Asn Ser Lys Asn Ala Glu Ser Phe
 2545 2550 2555 2560
 Gln Phe His Thr Glu Asp Tyr Phe Gly Lys Glu Ser Leu Trp Thr Gly
 2565 2570 2575
 Lys Gly Ile Gln Leu Ala Asp Gly Gly Trp Leu Ile Pro Ser Asn Asp
 2580 2585 2590
 Gly Lys Ala Gly Lys Glu Glu Phe Tyr Arg Ala Leu Cys Asp Thr Pro
 2595 2600 2605
 Gly Val Asp Pro Lys Leu Ile Ser Arg Ile Trp Val Tyr Asn His Tyr
 2610 2615 2620
 Arg Trp Ile Ile Trp Lys Leu Ala Ala Met Glu Cys Ala Phe Pro Lys
 2625 2630 2635 2640
 Glu Phe Ala Asn Arg Cys Leu Ser Pro Glu Arg Val Leu Leu Gln Leu
 2645 2650 2655
 Lys Tyr Arg Tyr Asp Thr Glu Ile Asp Arg Ser Arg Arg Ser Ala Ile
 2660 2665 2670
 Lys Lys Ile Met Glu Arg Asp Asp Thr Ala Ala Lys Thr Leu Val Leu
 2675 2680 2685
 Cys Val Ser Asp Ile Ile Ser Leu Ser Ala Asn Ile Ser Glu Thr Ser
 2690 2695 2700
 Ser Asn Lys Thr Ser Ser Ala Asp Thr Gln Lys Val Ala Ile Ile Glu
 2705 2710 2715 2720
 Leu Thr Asp Gly Trp Tyr Ala Val Lys Ala Gln Leu Asp Pro Pro Leu
 2725 2730 2735
 Leu Ala Val Leu Lys Asn Gly Arg Leu Thr Val Gly Gln Lys Ile Ile
 2740 2745 2750
 Leu His Gly Ala Glu Leu Val Gly Ser Pro Asp Ala Cys Thr Pro Leu
 2755 2760 2765
 Glu Ala Pro Glu Ser Leu Met Leu Lys Ile Ser Ala Asn Ser Thr Arg
 2770 2775 2780

Pro Ala Arg Trp Tyr Thr Lys Leu Gly Phe Phe Pro Asp Pro Arg Pro
 2785 2790 2795 2800
 Phe Pro Leu Pro Leu Ser Ser Leu Phe Ser Asp Gly Gly Asn Val GLY
 2805 2810 2815
 Cys Val Asp Val Ile Ile Gln Arg Ala Tyr Pro Ile Gln Trp Met Glu
 2820 2825 2830
 Lys Thr Ser Ser Gly Leu Tyr Ile Phe Arg Asn Glu Arg Glu Glu Glu
 2835 2840 2845
 Lys Glu Ala Ala Lys Tyr Val Glu Ala Gln Gln Lys Arg Leu Glu Ala
 2850 2855 2860
 Leu Phe Thr Lys Ile Gln Glu Glu Phe Glu His Glu Glu Asn Thr
 2865 2870 2875 2880
 Thr Lys Pro Tyr Leu Pro Ser Arg Ala Leu Thr Arg Gln Gln Val Arg
 2885 2890 2895
 Ala Leu Gln Asp Gly Ala Glu Leu Tyr Glu Ala Val Lys Asn Ala Ala
 2900 2905 2910
 Asp Pro Ala Tyr Leu Glu Gly Tyr Phe Ser Glu Glu Gln Leu Arg Ala
 2915 2920 2925
 Leu Asn Asn His Arg Gln Met Leu Asn Asp Lys Lys Gln Ala Gln Ile
 2930 2935 2940
 Gln Leu Glu Ile Arg Lys Ala Met Glu Ser Ala Glu Gln Lys Glu Gln
 2945 2950 2955 2960
 Gly Leu Ser Arg Asp Val Thr Thr Val Trp Lys Leu Arg Ile Val Ser
 2965 2970 2975
 Tyr Ser Lys Lys Glu Lys Asp Ser Val Ile Leu Ser Ile Trp Arg Pro
 2980 2985 2990
 Ser Ser Asp Leu Tyr Ser Leu Leu Thr Glu Gly Lys Arg Tyr Arg Ile
 2995 3000 3005
 Tyr His Leu Ala Thr Ser Lys Ser Lys Ser Glu Arg Ala Asn
 3010 3015 3020
 Ile Gln Leu Ala Ala Thr Lys Lys Thr Gln Tyr Gln Gln Leu Pro Val
 3025 3030 3035 3040
 Ser Asp Glu Ile Leu Phe Gln Ile Tyr Gln Pro Arg Glu Pro Leu His
 3045 3050 3055
 Phe Ser Lys Phe Leu Asp Pro Asp Phe Gln Pro Ser Cys Ser Glu Val
 3060 3065 3070
 Asp Leu Ile Gly Phe Val Val Ser Val Val Lys Lys Thr Gly Leu Ala
 3075 3080 3085
 Pro Phe Val Tyr Leu Ser Asp Glu Cys Tyr Asn Leu Leu Ala Ile Lys
 3090 3095 3100
 Phe Trp Ile Asp Leu Asn Glu Asp Ile Ile Lys Pro His Met Leu Ile
 3105 3110 3115 3120
 Ala Ala Ser Asn Leu Gln Trp Arg Pro Glu Ser Lys Ser Gly Leu Leu
 3125 3130 3135
 Thr Leu Phe Ala Gly Asp Phe Ser Val Phe Ser Ala Ser Pro Lys Glu
 3140 3145 3150
 Gly His Phe Gln Glu Thr Phe Asn Lys Met Lys Asn Thr Val Glu Asn
 3155 3160 3165
 Ile Asp Ile Leu Cys Asn Glu Ala Glu Asn Lys Leu Met His Ile Leu
 3170 3175 3180
 His Ala Asn Asp Pro Lys Trp Ser Thr Pro Thr Lys Asp Cys Thr Ser
 3185 3190 3195 3200
 Gly Pro Tyr Thr Ala Gln Ile Ile Pro Gly Thr Gly Asn Lys Leu Leu
 3205 3210 3215
 Met Ser Ser Pro Asn Cys Glu Ile Tyr Tyr Gln Ser Pro Leu Ser Leu
 3220 3225 3230
 Cys Met Ala Lys Arg Lys Ser Val Ser Thr Pro Val Ser Ala Gln Met
 3235 3240 3245
 Thr Ser Lys Ser Cys Lys Gly Glu Lys Glu Ile Asp Asp Gln Lys Asn
 3250 3255 3260
 Cys Lys Lys Arg Arg Ala Leu Asp Phe Leu Ser Arg Leu Pro Leu Pro

3265 3270 3275 3280
 Pro Pro Val Ser Pro Ile Cys Thr Phe Val Ser Pro Ala Ala Gln Lys
 3285 3290 3295
 Ala Phe Gln Pro Pro Arg Ser Cys Gly Thr Lys Tyr Glu Thr Pro Ile
 3300 3305 3310
 Lys Lys Lys Glu Leu Asn Ser Pro Gln Met Thr Pro Phe Lys Lys Phe
 3315 3320 3325
 Asn Glu Ile Ser Leu Leu Glu Ser Asn Ser Ile Ala Asp Glu Glu Leu
 3330 3335 3340
 Ala Leu Ile Asn Thr Gln Ala Leu Leu Ser Gly Ser Thr Gly Glu Lys
 3345 3350 3355 3360
 Gln Phe Ile Ser Val Ser Glu Ser Thr Arg Thr Ala Pro Thr Ser Ser
 3365 3370 3375
 Glu Asp Tyr Leu Arg Leu Lys Arg Arg Cys Thr Thr Ser Leu Ile Lys
 3380 3385 3390
 Glu Gln Glu Ser Ser Gln Ala Ser Thr Glu Glu Cys Glu Lys Asn Lys
 3395 3400 3405
 Gln Asp Thr Ile Thr Thr Lys Lys Tyr Ile
 3410 3415

(2) INFORMATION FOR SEQ ID NO:10:

D
D
COM
(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10485 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 229...10482
- (D) OTHER INFORMATION: BRCA2 (OMI4)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

| | | | | | | |
|---|-------------|-------------|-------------|-------------|-------------|-----|
| GGTGGCGCGA | GCTTCTGAAA | CTAGGCGGCA | GAGGC GGAGC | CGCTGTGGCA | CTGCTGCGCC | 60 |
| TCTGCTGCGC | CTCGGGGTGTC | TTTGCGGGCG | GTGGGTCGCC | GCCGGGAGAA | GCGTGAGGGGG | 120 |
| ACAGATTGTTGT | GACCGGGCGCG | GT TTT GTCA | GCTTACTCCG | GCCAAAAAAAG | AACTGCACCT | 180 |
| CTGGAGCGGA | CTTATTTACC | AAGCATTGGA | GGAATATCGT | AGGTAAAAA | ATG CCT ATT | 237 |
| | | | | | Met Pro Ile | |
| | | | | | 1 | |
| GGA TCC AAA GAG AGG CCA ACA TTT TTT GAA ATT TTT AAG ACA CGC TGC | | | | | | 285 |
| Gly Ser Lys Glu Arg Pro Thr Phe Phe Glu Ile Phe Lys Thr Arg Cys | | | | | | |
| 5 | 10 | 15 | | | | |
| AAC AAA GCA GAT TTA GGA CCA ATA AGT CTT AAT TGG TTT GAA GAA CTT | | | | | | 333 |
| Asn Lys Ala Asp Leu Gly Pro Ile Ser Leu Asn Trp Phe Glu Glu Leu | | | | | | |
| 20 | 25 | 30 | 35 | | | |
| TCT TCA GAA GCT CCA CCC TAT AAT TCT GAA CCT GCA GAA GAA TCT GAA | | | | | | 381 |
| Ser Ser Glu Ala Pro Pro Tyr Asn Ser Glu Pro Ala Glu Glu Ser Glu | | | | | | |
| 40 | 45 | 50 | | | | |
| CAT AAA AAC AAC AAT TAC GAA CCA AAC CTA TTT AAA ACT CCA CAA AGG | | | | | | 429 |
| His Lys Asn Asn Asn Tyr Glu Pro Asn Leu Phe Lys Thr Pro Gln Arg | | | | | | |
| 55 | 60 | 65 | | | | |

AAA CCA TCT TAT AAT CAG CTG GCT TCA ACT CCA ATA ATA TTC AAA GAG 477
 Lys Pro Ser Tyr Asn Gln Leu Ala Ser Thr Pro Ile Ile Phe Lys Glu
 70 75 80

CAA GGG CTG ACT CTG CCG CTG TAC CAA TCT CCT GTA AAA GAA TTA GAT 525
 Gln Gly Leu Thr Leu Pro Leu Tyr Gln Ser Pro Val Lys Glu Leu Asp
 85 90 95

AAA TTC AAA TTA GAC TTA GGA AGG AAT GTT CCC AAT AGT AGA CAT AAA 573
 Lys Phe Lys Leu Asp Leu Gly Arg Asn Val Pro Asn Ser Arg His Lys
 100 105 110 115

AGT CTT CGC ACA GTG AAA ACT AAA ATG GAT CAA GCA GAT GAT GTT TCC 621
 Ser Leu Arg Thr Val Lys Thr Lys Met Asp Gln Ala Asp Asp Val Ser
 120 125 130

TGT CCA CTT CTA AAT TCT TGT CTT AGT GAA AGT CCT GTT GTT CTA CAA 669
 Cys Pro Leu Leu Asn Ser Cys Leu Ser Glu Ser Pro Val Val Leu Gln
 135 140 145

TGT ACA CAT GTA ACA CCA CAA AGA GAT AAG TCA GTG GTA TGT GGG AGT 717
 Cys Thr His Val Thr Pro Gln Arg Asp Lys Ser Val Val Cys Gly Ser
 150 155 160

TTG TTT CAT ACA CCA AAG TTT GTG AAG GGT CGT CAG ACA CCA AAA CAT 765
 Leu Phe His Thr Pro Lys Phe Val Lys Gly Arg Gln Thr Pro Lys His
 165 170 175

ATT TCT GAA AGT CTA GGA GCT GAG GTG GAT CCT GAT ATG TCT TGG TCA 813
 Ile Ser Glu Ser Leu Gly Ala Glu Val Asp Pro Asp Met Ser Trp Ser
 180 185 190 195

AGT TCT TTA GCT ACA CCA CCC ACC CTT AGT TCT ACT GTG CTC ATA GTC 861
 Ser Ser Leu Ala Thr Pro Pro Thr Leu Ser Ser Thr Val Leu Ile Val
 200 205 210

AGA AAT GAA GAA GCA TCT GAA ACT GTA TTT CCT CAT GAT ACT ACT GCT 909
 Arg Asn Glu Glu Ala Ser Glu Thr Val Phe Pro His Asp Thr Thr Ala
 215 220 225

AAT GTG AAA AGC TAT TTT TCC AAT CAT GAT GAA AGT CTG AAG AAA AAT 957
 Asn Val Lys Ser Tyr Phe Ser Asn His Asp Glu Ser Leu Lys Lys Asn
 230 235 240

GAT AGA TTT ATC GCT TCT GTG ACA GAC AGT GAA AAC ACA AAT CAA AGA 1005
 Asp Arg Phe Ile Ala Ser Val Thr Asp Ser Glu Asn Thr Asn Gln Arg
 245 250 255

GAA GCT GCA AGT CAT GGA TTT GGA AAA ACA TCA GGG AAT TCA TTT AAA 1053
 Glu Ala Ala Ser His Gly Phe Gly Lys Thr Ser Gly Asn Ser Phe Lys
 260 265 270 275

GTA AAT AGC TGC AAA GAC CAC ATT GGA AAG TCA ATG CCA AAT GTC CTA 1101
 Val Asn Ser Cys Lys Asp His Ile Gly Lys Ser Met Pro Asn Val Leu
 280 285 290

GAA GAT GAA GTA TAT GAA ACA GTT GTA GAT ACC TCT GAA GAA GAT AGT 1149
 Glu Asp Glu Val Tyr Glu Thr Val Val Asp Thr Ser Glu Glu Asp Ser
 295 300 305

TTT TCA TTA TGT TTT TCT AAA TGT AGA ACA AAA AAT CTA CAA AAA GTA 1197

| | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|------|
| Phe | Ser | Leu | Cys | Phe | Ser | Lys | Cys | Arg | Thr | Lys | Asn | Leu | Gln | Lys | Val | |
| | | | | | | | | | | | | | | | | 310 |
| | | | | | | | | | | | | | | | | 315 |
| | | | | | | | | | | | | | | | | 320 |
| AGA | ACT | AGC | AAG | ACT | AGG | AAA | AAA | ATT | TTC | CAT | GAA | GCA | AAC | GCT | GAT | 1245 |
| Arg | Thr | Ser | Lys | Thr | Arg | Lys | Lys | Ile | Phe | His | Glu | Ala | Asn | Ala | Asp | |
| | | | | | | | | | | | | | | | | 325 |
| | | | | | | | | | | | | | | | | 330 |
| | | | | | | | | | | | | | | | | 335 |
| GAA | TGT | GAA | AAA | TCT | AAA | AAC | CAA | GTG | AAA | GAA | AAA | TAC | TCA | TTT | GTA | 1293 |
| Glu | Cys | Glu | Lys | Ser | Lys | Asn | Gln | Val | Lys | Glu | Lys | Tyr | Ser | Phe | Val | |
| | | | | | | | | | | | | | | | | 340 |
| | | | | | | | | | | | | | | | | 345 |
| | | | | | | | | | | | | | | | | 350 |
| | | | | | | | | | | | | | | | | 355 |
| TCT | GAA | GTG | GAA | CCA | AAT | GAT | ACT | GAT | CCA | TTA | GAT | TCA | AAT | GTA | GCA | 1341 |
| Ser | Glu | Val | Glu | Pro | Asn | Asp | Thr | Asp | Pro | Leu | Asp | Ser | Asn | Val | Ala | |
| | | | | | | | | | | | | | | | | 360 |
| | | | | | | | | | | | | | | | | 365 |
| | | | | | | | | | | | | | | | | 370 |
| CAT | CAG | AAG | CCC | TTT | GAG | AGT | GGA | AGT | GAC | AAA | ATC | TCC | AAG | GAA | GTT | 1389 |
| His | Gln | Lys | Pro | Phe | Glu | Ser | Gly | Ser | Asp | Lys | Ile | Ser | Lys | Glu | Val | |
| | | | | | | | | | | | | | | | | 375 |
| | | | | | | | | | | | | | | | | 380 |
| | | | | | | | | | | | | | | | | 385 |
| GTA | CCG | TCT | TTG | GCC | TGT | GAA | TGG | TCT | CAA | CTA | ACC | CTT | TCA | GGT | CTA | 1437 |
| Val | Pro | Ser | Leu | Ala | Cys | Glu | Trp | Ser | Gln | Leu | Thr | Leu | Ser | Gly | Leu | |
| | | | | | | | | | | | | | | | | 390 |
| | | | | | | | | | | | | | | | | 395 |
| | | | | | | | | | | | | | | | | 400 |
| AAT | GGA | GCC | CAG | ATG | GAG | AAA | ATA | CCC | CTA | TTG | CAT | ATT | TCT | TCA | TGT | 1485 |
| Asn | Gly | Ala | Gln | Met | Glu | Lys | Ile | Pro | Leu | Leu | His | Ile | Ser | Ser | Cys | |
| | | | | | | | | | | | | | | | | 405 |
| | | | | | | | | | | | | | | | | 410 |
| | | | | | | | | | | | | | | | | 415 |
| GAC | CAA | AAT | ATT | TCA | GAA | AAA | GAC | CTA | TTA | GAC | ACA | GAG | AAC | AAA | AGA | 1533 |
| Asp | Gln | Asn | Ile | Ser | Glu | Lys | Asp | Leu | Leu | Asp | Thr | Glu | Asn | Lys | Arg | |
| | | | | | | | | | | | | | | | | 420 |
| | | | | | | | | | | | | | | | | 425 |
| | | | | | | | | | | | | | | | | 430 |
| | | | | | | | | | | | | | | | | 435 |
| AAG | AAA | GAT | TTT | CTT | ACT | TCA | GAG | AAT | TCT | TTG | CCA | CGT | ATT | TCT | AGC | 1581 |
| Lys | Lys | Asp | Phe | Leu | Thr | Ser | Glu | Asn | Ser | Leu | Pro | Arg | Ile | Ser | Ser | |
| | | | | | | | | | | | | | | | | 440 |
| | | | | | | | | | | | | | | | | 445 |
| | | | | | | | | | | | | | | | | 450 |
| CTA | CCA | AAA | TCA | GAG | AAG | CCA | TTA | AAT | GAG | GAA | ACA | GTG | GTA | AAT | AAG | 1629 |
| Leu | Pro | Lys | Ser | Glu | Lys | Pro | Leu | Asn | Glu | Glu | Thr | Val | Val | Asn | Lys | |
| | | | | | | | | | | | | | | | | 455 |
| | | | | | | | | | | | | | | | | 460 |
| | | | | | | | | | | | | | | | | 465 |
| AGA | GAT | GAA | GAG | CAG | CAT | CTT | GAA | TCT | CAT | ACA | GAC | TGC | ATT | CTT | GCA | 1677 |
| Arg | Asp | Glu | Glu | Gln | His | Leu | Glu | Ser | His | Thr | Asp | Cys | Ile | Leu | Ala | |
| | | | | | | | | | | | | | | | | 470 |
| | | | | | | | | | | | | | | | | 475 |
| | | | | | | | | | | | | | | | | 480 |
| GTA | AAG | CAG | GCA | ATA | TCT | GGA | ACT | TCT | CCA | GTG | GCT | TCT | TCA | TTT | CAG | 1725 |
| Val | Lys | Gln | Ala | Ile | Ser | Gly | Thr | Ser | Pro | Val | Ala | Ser | Ser | Phe | Gln | |
| | | | | | | | | | | | | | | | | 485 |
| | | | | | | | | | | | | | | | | 490 |
| | | | | | | | | | | | | | | | | 495 |
| GGT | ATC | AAA | AAG | TCT | ATA | TTC | AGA | ATA | AGA | GAA | TCA | CCT | AAA | GAG | ACT | 1773 |
| Gly | Ile | Lys | Lys | Ser | Ile | Phe | Arg | Ile | Arg | Glu | Ser | Pro | Lys | Glu | Thr | |
| | | | | | | | | | | | | | | | | 500 |
| | | | | | | | | | | | | | | | | 505 |
| | | | | | | | | | | | | | | | | 510 |
| | | | | | | | | | | | | | | | | 515 |
| TTC | AAT | GCA | AGT | TTT | TCA | GGT | CAT | ATG | ACT | GAT | CCA | AAC | TTT | AAA | AAA | 1821 |
| Phe | Asn | Ala | Ser | Phe | Ser | Gly | His | Met | Thr | Asp | Pro | Asn | Phe | Lys | Lys | |
| | | | | | | | | | | | | | | | | 520 |
| | | | | | | | | | | | | | | | | 525 |
| | | | | | | | | | | | | | | | | 530 |
| GAA | ACT | GAA | GCC | TCT | GAA | AGT | GGA | CTG | GAA | ATA | CAT | ACT | GTT | TGC | TCA | 1869 |
| Glu | Thr | Glu | Ala | Ser | Glu | Ser | Gly | Leu | Glu | Ile | His | Thr | Val | Cys | Ser | |
| | | | | | | | | | | | | | | | | 535 |
| | | | | | | | | | | | | | | | | 540 |
| | | | | | | | | | | | | | | | | 545 |
| CAG | AAG | GAG | GAC | TCC | TTA | TGT | CCA | AAT | TTA | ATT | GAT | AAT | GGA | AGC | TGG | 1917 |
| Gln | Lys | Glu | Asp | Ser | Leu | Cys | Pro | Asn | Leu | Ile | Asp | Asn | Gly | Ser | Trp | |

| 550 | 555 | 560 | |
|---|-----|-----|------|
| CCA GCC ACC ACC ACA CAG AAT TCT GTA GCT TTG AAG AAT GCA GGT TTA Pro Ala Thr Thr Thr Gln Asn Ser Val Ala Leu Lys Asn Ala Gly Leu 565 | 570 | 575 | 1965 |
| ATA TCC ACT TTG AAA AAG AAA ACA AAT AAG TTT ATT TAT GCT ATA CAT Ile Ser Thr Leu Lys Lys Thr Asn Lys Phe Ile Tyr Ala Ile His 580 | 585 | 590 | 2013 |
| GAT GAA ACA TCT TAT AAA GGA AAA AAA ATA CCG AAA GAC CAA AAA TCA Asp Glu Thr Ser Tyr Lys Gly Lys Lys Ile Pro Lys Asp Gln Lys Ser 600 | 605 | 610 | 2061 |
| GAA CTA ATT AAC TGT TCA GCC CAG TTT GAA GCA AAT GCT TTT GAA GCA Glu Leu Ile Asn Cys Ser Ala Gln Phe Glu Ala Asn Ala Phe Glu Ala 615 | 620 | 625 | 2109 |
| D W K | | | |
| CCA CTT ACA TTT GCA AAT GCT GAT TCA GGT TTA TTG CAT TCT TCT GTG Pro Leu Thr Phe Ala Asn Ala Asp Ser Gly Leu Leu His Ser Ser Val 630 | 635 | 640 | 2157 |
| AAA AGA AGC TGT TCA CAG AAT GAT TCT GAA GAA CCA ACT TTG TCC TTA Lys Arg Ser Cys Ser Gln Asn Asp Ser Glu Glu Pro Thr Leu Ser Leu 645 | 650 | 655 | 2205 |
| ACT AGC TCT TTT GGG ACA ATT CTG AGG AAA TGT TCT AGA AAT GAA ACA Thr Ser Ser Phe Gly Thr Ile Leu Arg Lys Cys Ser Arg Asn Glu Thr 660 | 665 | 670 | 2253 |
| TGT TCT AAT ACA GTA ATC TCT CAG GAT CTT GAT TAT AAA GAA GCA Cys Ser Asn Asn Thr Val Ile Ser Gln Asp Leu Asp Tyr Lys Glu Ala 680 | 685 | 690 | 2301 |
| AAA TGT AAT AAG GAA AAA CTA CAG TTA TTT ATT ACC CCA GAA GCT GAT Lys Cys Asn Lys Glu Lys Leu Gln Leu Phe Ile Thr Pro Glu Ala Asp 695 | 700 | 705 | 2349 |
| TCT CTG TCA TGC CTG CAG GAA GGA CAG TGT GAA AAT GAT CCA AAA AGC Ser Leu Ser Cys Leu Gln Glu Gly Gln Cys Glu Asn Asp Pro Lys Ser 710 | 715 | 720 | 2397 |
| AAA AAA GTT TCA GAT ATA AAA GAA GAG GTC TTG GCT GCA GCA TGT CAC Lys Lys Val Ser Asp Ile Lys Glu Glu Val Leu Ala Ala Ala Cys His 725 | 730 | 735 | 2445 |
| CCA GTA CAA CAT TCA AAA GTG GAA TAC AGT GAT ACT GAC TTT CAA TCC Pro Val Gln His Ser Lys Val Glu Tyr Ser Asp Thr Asp Phe Gln Ser 740 | 745 | 750 | 2493 |
| CAG AAA AGT CTT TTA TAT GAT CAT GAA AAT GCC AGC ACT CTT ATT TTA Gln Lys Ser Leu Leu Tyr Asp His Glu Asn Ala Ser Thr Leu Ile Leu 760 | 765 | 770 | 2541 |
| ACT CCT ACT TCC AAG GAT GTT CTG TCA AAC CTA GTC ATG ATT TCT AGA Thr Pro Thr Ser Lys Asp Val Leu Ser Asn Leu Val Met Ile Ser Arg 775 | 780 | 785 | 2589 |
| GGC AAA GAA TCA TAC AAA ATG TCA GAC AAG CTC AAA GGT AAC AAT TAT Gly Lys Glu Ser Tyr Lys Met Ser Asp Lys Leu Lys Gly Asn Asn Tyr 790 | 795 | 800 | 2637 |

GAA TCT GAT GTT GAA TTA ACC AAA AAT ATT CCC ATG GAA AAG AAT CAA 2685
 Glu Ser Asp Val Glu Leu Thr Lys Asn Ile Pro Met Glu Lys Asn Gln
 805 810 815

GAT GTA TGT GCT TTA AAT GAA AAT TAT AAA AAC GTT GAG CTG TTG CCA 2733
 Asp Val Cys Ala Leu Asn Glu Asn Tyr Lys Asn Val Glu Leu Leu Pro
 820 825 830 835

CCT GAA AAA TAC ATG AGA GTA GCA TCA CCT TCA AGA AAG GTA CAA TTC 2781
 Pro Glu Lys Tyr Met Arg Val Ala Ser Pro Ser Arg Lys Val Gln Phe
 840 845 850

AAC CAA AAC ACA AAT CTA AGA GTA ATC CAA AAA AAT CAA GAA GAA ACT 2829
 Asn Gln Asn Thr Asn Leu Arg Val Ile Gln Lys Asn Gln Glu Glu Thr
 855 860 865

ACT TCA ATT TCA AAA ATA ACT GTC AAT CCA GAC TCT GAA GAA CTT TTC 2877
 Thr Ser Ile Ser Lys Ile Thr Val Asn Pro Asp Ser Glu Glu Leu Phe
 870 875 880

*D /
cont*
 TCA GAC AAT GAG AAT AAT TTT GTC TTC CAA GTA GCT AAT GAA AGG AAT 2925
 Ser Asp Asn Glu Asn Asn Phe Val Phe Gln Val Ala Asn Glu Arg Asn
 885 890 895

AAT CTT GCT TTA GGA AAT ACT AAG GAA CTT CAT GAA ACA GAC TTG ACT 2973
 Asn Leu Ala Leu Gly Asn Thr Lys Glu Leu His Glu Thr Asp Leu Thr
 900 905 910 915

TGT GTA AAC GAA CCC ATT TTC AAG AAC TCT ACC ATG GTT TTA TAT GGA 3021
 Cys Val Asn Glu Pro Ile Phe Lys Asn Ser Thr Met Val Leu Tyr Gly
 920 925 930

GAC ACA GGT GAT AAA CAA GCA ACC CAA GTG TCA ATT AAA AAA GAT TTG 3069
 Asp Thr Gly Asp Lys Gln Ala Thr Gln Val Ser Ile Lys Lys Asp Leu
 935 940 945

GTT TAT GTT CTT GCA GAG GAG AAC AAA AAT AGT GTA AAG CAG CAT ATA 3117
 Val Tyr Val Leu Ala Glu Glu Asn Lys Asn Ser Val Lys Gln His Ile
 950 955 960

AAA ATG ACT CTA GGT CAA GAT TTA AAA TCG GAC ATC TCC TTG AAT ATA 3165
 Lys Met Thr Leu Gly Gln Asp Leu Lys Ser Asp Ile Ser Leu Asn Ile
 965 970 975

GAT AAA ATA CCA GAA AAA AAT AAT GAT TAC ATG AAC AAA TGG GCA GGA 3213
 Asp Lys Ile Pro Glu Lys Asn Asn Asp Tyr Met Asn Lys Trp Ala Gly
 980 985 990 995

CTC TTA GGT CCA ATT TCA AAT CAC AGT TTT GGA GGT AGC TTC AGA ACA 3261
 Leu Leu Gly Pro Ile Ser Asn His Ser Phe Gly Gly Ser Phe Arg Thr
 1000 1005 1010

GCT TCA AAT AAG GAA ATC AAG CTC TCT GAA CAT AAC ATT AAG AAG AGC 3309
 Ala Ser Asn Lys Glu Ile Lys Leu Ser Glu His Asn Ile Lys Lys Ser
 1015 1020 1025

AAA ATG TTC TTC AAA GAT ATT GAA GAA CAA TAT CCT ACT AGT TTA GCT 3357
 Lys Met Phe Phe Lys Asp Ile Glu Glu Gln Tyr Pro Thr Ser Leu Ala
 1030 1035 1040

D
cont

| | |
|---|------|
| TGT GTT GAA ATT GTA AAT ACC TTG GCA TTA GAT AAT CAA AAG AAA CTG Cys Val Glu Ile Val Asn Thr Leu Ala Leu Asp Asn Gln Lys Lys Leu 1045 1050 1055 | 3405 |
| AGC AAG CCT CAG TCA ATT AAT ACT GTA TCT GCA CAT TTA CAG AGT AGT Ser Lys Pro Gln Ser Ile Asn Thr Val Ser Ala His Leu Gln Ser Ser 1060 1065 1070 1075 | 3453 |
| GTA GTT GTT TCT GAT TGT AAA AAT AGT CAT ATA ACC CCT CAG ATG TTA Val Val Val Ser Asp Cys Lys Asn Ser His Ile Thr Pro Gln Met Leu 1080 1085 1090 | 3501 |
| TTT TCC AAG CAG GAT TTT AAT TCA AAC CAT AAT TTA ACA CCT AGC CAA Phe Ser Lys Gln Asp Phe Asn Ser His Asn Leu Thr Pro Ser Gln 1095 1100 1105 | 3549 |
| AAG GCA GAA ATT ACA GAA CTT TCT ACT ATA TTA GAA GAA TCA CGA AGT Lys Ala Glu Ile Thr Glu Leu Ser Thr Ile Leu Glu Ser Gly Ser 1110 1115 1120 | 3597 |
| CAG TTT GAA TTT ACT CAG TTT AGA AAG CCA AGC TAC ATA TTG CAG AAG Gln Phe Glu Phe Thr Gln Phe Arg Lys Pro Ser Tyr Ile Leu Gln Lys 1125 1130 1135 | 3645 |
| AGT ACA TTT GAA GTG CCT GAA AAC CAG ATG ACT ATC TTA AAG ACC ACT Ser Thr Phe Glu Val Pro Glu Asn Gln Met Thr Ile Leu Lys Thr Thr 1140 1145 1150 1155 | 3693 |
| TCT GAG GAA TGC AGA GAT GCT GAT CTT CAT GTC ATA ATG AAT GCC CCA Ser Glu Glu Cys Arg Asp Ala Asp Leu His Val Ile Met Asn Ala Pro 1160 1165 1170 | 3741 |
| TCG ATT GGT CAG GTA GAC AGC AGC AAG CAA TTT GAA GGT ACA GTT GAA Ser Ile Gly Gln Val Asp Ser Ser Lys Gln Phe Glu Gly Thr Val Glu 1175 1180 1185 | 3789 |
| ATT AAA CGG AAG TTT GCT GGC CTG TTG AAA AAT GAC TGT AAC AAA ACT Ile Lys Arg Lys Phe Ala Gly Leu Leu Lys Asn Asp Cys Asn Lys Ser 1190 1195 1200 | 3837 |
| GCT TCT GGT TAT TTA ACA GAT GAA AAT GAA GTG GGG TTT AGG GGC TTT Ala Ser Gly Tyr Leu Thr Asp Glu Asn Glu Val Gly Phe Arg Gly Phe 1205 1210 1215 | 3885 |
| TAT TCT GCT CAT GGC ACA AAA CTG AAT GTT TCT ACT GAA GCT CTG CAA Tyr Ser Ala His Gly Thr Lys Leu Asn Val Ser Thr Glu Ala Leu Gln 1220 1225 1230 1235 | 3933 |
| AAA GCT GTG AAA CTG TTT AGT GAT ATT GAG AAT ATT AGT GAG GAA ACT Lys Ala Val Lys Leu Phe Ser Asp Ile Glu Asn Ile Ser Glu Glu Thr 1240 1245 1250 | 3981 |
| TCT GCA GAG GTA CAT CCA ATA AGT TTA TCT TCA AGT AAA TGT CAT GAT Ser Ala Glu Val His Pro Ile Ser Leu Ser Ser Ser Lys Cys His Asp 1255 1260 1265 | 4029 |
| TCT GTT GTT TCA ATG TTT AAG ATA GAA AAT CAT AAT GAT AAA ACT GTA Ser Val Val Ser Met Phe Lys Ile Glu Asn His Asn Asp Lys Thr Val 1270 1275 1280 | 4077 |
| AGT GAA AAA AAT AAT AAA TGC CAA CTG ATA TTA CAA AAT AAT ATT GAA | 4125 |

Ser Glu Lys Asn Asn Lys Cys Gln Leu Ile Leu Gln Asn Asn Ile Glu
 1285 1290 1295

ATG ACT ACT GGC ACT TTT GTT GAA GAA ATT ACT GAA AAT TAC AAG AGA 4173
 Met Thr Thr Gly Thr Phe Val Glu Glu Ile Thr Glu Asn Tyr Lys Arg
 1300 1305 1310 1315

AAT ACT GAA AAT GAA GAT AAC AAA TAT ACT GCT GCC AGT AGA AAT TCT 4221
 Asn Thr Glu Asn Glu Asp Asn Lys Tyr Thr Ala Ala Ser Arg Asn Ser
 1320 1325 1330

CAT AAC TTA GAA TTT GAT GGC AGT GAT TCA AGT AAA AAT GAT ACT GTT 4269
 His Asn Leu Glu Phe Asp Gly Ser Asp Ser Ser Lys Asn Asp Thr Val
 1335 1340 1345

TGT ATT CAT AAA GAT GAA ACG GAC TTG CTA TTT ACT GAT CAG CAC AAC 4317
 Cys Ile His Lys Asp Glu Thr Asp Leu Leu Phe Thr Asp Gln His Asn
 1350 1355 1360

ATA TGT CTT AAA TTA TCT GGC CAG TTT ATG AAG GAG GGA AAC ACT CAG 4365
 Ile Cys Leu Lys Leu Ser Gly Gln Phe Met Lys Glu Gly Asn Thr Gln
 1365 1370 1375

ATT AAA GAA GAT TTG TCA GAT TTA ACT TTT TTG GAA GTT GCG AAA GCT 4413
 Ile Lys Glu Asp Leu Ser Asp Leu Thr Phe Leu Glu Val Ala Lys Ala
 1380 1385 1390 1395

CAA GAA GCA TGT CAT GGT AAT ACT TCA AAT AAA GAA CAG TTA ACT GCT 4461
 Gln Glu Ala Cys His Gly Asn Thr Ser Asn Lys Glu Gln Leu Thr Ala
 1400 1405 1410

ACT AAA ACG GAG CAA AAT ATA AAA GAT TTT GAG ACT TCT GAT ACA TTT 4509
 Thr Lys Thr Glu Gln Asn Ile Lys Asp Phe Glu Thr Ser Asp Thr Phe
 1415 1420 1425

TTT CAG ACT GCA AGT GGG AAA AAT ATT AGT GTC GCC AAA GAG TCA TTT 4557
 Phe Gln Thr Ala Ser Gly Lys Asn Ile Ser Val Ala Lys Glu Ser Phe
 1430 1435 1440

AAT AAA ATT GTA AAT TTC TTT GAT CAG AAA CCA GAA GAA TTG CAT AAC 4605
 Asn Lys Ile Val Asn Phe Phe Asp Gln Lys Pro Glu Glu Leu His Asn
 1445 1450 1455

TTT TCC TTA AAT TCT GAA TTA CAT TCT GAC ATA AGA AAG AAC AAA ATG 4653
 Phe Ser Leu Asn Ser Glu Leu His Ser Asp Ile Arg Lys Asn Lys Met
 1460 1465 1470 1475

GAC ATT CTA AGT TAT GAG GAA ACA GAC ATA GTT AAA CAC AAA ATA CTG 4701
 Asp Ile Leu Ser Tyr Glu Glu Thr Asp Ile Val Lys His Lys Ile Leu
 1480 1485 1490

AAA GAA AGT GTC CCA GTT GGT ACT GGA AAT CAA CTA GTG ACC TTC CAG 4749
 Lys Glu Ser Val Pro Val Gly Thr Gly Asn Gln Leu Val Thr Phe Gln
 1495 1500 1505

GGA CAA CCC GAA CGT GAT GAA AAG ATC AAA GAA CCT ACT CTG TTG GGT 4797
 Gly Gln Pro Glu Arg Asp Glu Lys Ile Lys Glu Pro Thr Leu Leu Gly
 1510 1515 1520

TTT CAT ACA GCT AGC GGG AAA AAA GTT AAA ATT GCA AAG GAA TCT TTG 4845
 Phe His Thr Ala Ser Gly Lys Lys Val Lys Ile Ala Lys Glu Ser Leu

1525 1530 1535 4893
 GAC AAA GTG AAA AAC CTT TTT GAT GAA AAA GAG CAA GGT ACT AGT GAA
 Asp Lys Val Lys Asn Leu Phe Asp Glu Lys Glu Gln Gly Thr Ser Glu
 1540 1545 1550 1555

 ATC ACC AGT TTT AGC CAT CAA TGG GCA AAG ACC CTA AAG TAC AGA GAG 4941
 Ile Thr Ser Phe Ser His Gln Trp Ala Lys Thr Leu Lys Tyr Arg Glu
 1560 1565 1570

 GCC TGT AAA GAC CTT GAA TTA GCA TGT GAG ACC ATT GAG ATC ACA GCT 4989
 Ala Cys Lys Asp Leu Glu Leu Ala Cys Glu Thr Ile Glu Ile Thr Ala
 1575 1580 1585

 GCC CCA AAG TGT AAA GAA ATG CAG AAT TCT CTC AAT AAT GAT AAA AAC 5037
 Ala Pro Lys Cys Lys Glu Met Gln Asn Ser Leu Asn Asn Asp Lys Asn
 1590 1595 1600

 CTT GTT TCT ATT GAG ACT GTG GTG CCA CCT AAG CTC ATT AGT GAT AAT 5085
 Leu Val Ser Ile Glu Thr Val Val Pro Pro Lys Leu Leu Ser Asp Asn
 1605 1610 1615


 TTA TGT AGA CAA ACT GAA AAT CTC AAA ACA TCA AAA AGT ATC TTT TTG 5133
 Leu Cys Arg Gln Thr Glu Asn Leu Lys Thr Ser Lys Ser Ile Phe Leu
 1620 1625 1630 1635

 AAA GTT AAA GTA CAT GAA AAT GTA GAA AAA GAA ACA GCA AAA AGT CCT 5181
 Lys Val Lys Val His Glu Asn Val Glu Lys Glu Thr Ala Lys Ser Pro
 1640 1645 1650

 GCA ACT TGT TAC ACA AAT CAG TCC CCT TAT TCA GTC ATT GAA AAT TCA 5229
 Ala Thr Cys Tyr Thr Asn Gln Ser Pro Tyr Ser Val Ile Glu Asn Ser
 1655 1660 1665

 GCC TTA GCT TTT TAC ACA AGT TGT AGT AGA AAA ACT TCT GTG AGT CAG 5277
 Ala Leu Ala Phe Tyr Thr Ser Cys Ser Arg Lys Thr Ser Val Ser Gln
 1670 1675 1680

 ACT TCA TTA CTT GAA GCA AAA AAA TGG CTT AGA GAA GGA ATA TTT GAT 5325
 Thr Ser Leu Leu Glu Ala Lys Lys Trp Leu Arg Glu Gly Ile Phe Asp
 1685 1690 1695

 GGT CAA CCA GAA AGA ATA AAT ACT GCA GAT TAT GTA GGA AAT TAT TTG 5373
 Gly Gln Pro Glu Arg Ile Asn Thr Ala Asp Tyr Val Gly Asn Tyr Leu
 1700 1705 1710 1715

 TAT GAA AAT AAT TCA AAC AGT ACT ATA GCT GAA AAT GAC AAA AAT CAT 5421
 Tyr Glu Asn Asn Ser Asn Ser Thr Ile Ala Glu Asn Asp Lys Asn His
 1720 1725 1730

 CTC TCC GAA AAA CAA GAT ACT TAT TTA AGT AAC AGT AGC ATG TCT AAC 5469
 Leu Ser Glu Lys Gln Asp Thr Tyr Leu Ser Asn Ser Ser Met Ser Asn
 1735 1740 1745

 AGC TAT TCC TAC CAT TCT GAT GAG GTA TAT AAT GAT TCA GGA TAT CTC 5517
 Ser Tyr Ser Tyr His Ser Asp Glu Val Tyr Asn Asp Ser Gly Tyr Leu
 1750 1755 1760

 TCA AAA AAT AAA CTT GAT TCT GGT ATT GAG CCA GTA TTG AAG AAT GTT 5565
 Ser Lys Asn Lys Leu Asp Ser Gly Ile Glu Pro Val Leu Lys Asn Val
 1765 1770 1775

GAA GAT CAA AAA AAC ACT AGT TTT TCC AAA GTA ATA TCC AAT GTA AAA 5613
 Glu Asp Gln Lys Asn Thr Ser Phe Ser Lys Val Ile Ser Asn Val Lys
 1780 1785 1790 1795

GAT GCA AAT GCA TAC CCA CAA ACT GTA AAT GAA GAT ATT TGC GTT GAG 5661
 Asp Ala Asn Ala Tyr Pro Gln Thr Val Asn Glu Asp Ile Cys Val Glu
 1800 1805 1810

GAA CTT GTG ACT AGC TCT TCA CCC TGC AAA AAT AAA AAT GCA GCC ATT 5709
 Glu Leu Val Thr Ser Ser Pro Cys Lys Asn Lys Asn Ala Ala Ile
 1815 1820 1825

AAA TTG TCC ATA TCT AAT AGT AAT AAT TTT GAG GTA GGG CCA CCT GCA 5757
 Lys Leu Ser Ile Ser Asn Ser Asn Asn Phe Glu Val Gly Pro Pro Ala
 1830 1835 1840

TTT AGG ATA GCC AGT GGT AAA ATC GTT TGT GTT TCA CAT GAA ACA ATT 5805
 Phe Arg Ile Ala Ser Gly Lys Ile Val Cys Val Ser His Glu Thr Ile
 1845 1850 1855

AAA AAA GTG AAA GAC ATA TTT ACA GAC AGT TTC AGT AAA GTA ATT AAG 5853
 Lys Lys Val Lys Asp Ile Phe Thr Asp Ser Phe Ser Lys Val Ile Lys
 1860 1865 1870 1875

GAA AAC AAC GAG AAT AAA TCA AAA ATT TGC CAA ACG AAA ATT ATG GCA 5901
 Glu Asn Asn Glu Asn Lys Ser Lys Ile Cys Gln Thr Lys Ile Met Ala
 1880 1885 1890

GGT TGT TAC GAG GCA TTG GAT GAT TCA GAG GAT ATT CTT CAT AAC TCT 5949
 Gly Cys Tyr Glu Ala Leu Asp Asp Ser Glu Asp Ile Leu His Asn Ser
 1895 1900 1905

CTA GAT AAT GAT GAA TGT AGC ACG CAT TCA CAT AAG GTT TTT GCT GAC 5997
 Leu Asp Asn Asp Glu Cys Ser Thr His Ser His Lys Val Phe Ala Asp
 1910 1915 1920

ATT CAG AGT GAA GAA ATT TTA CAA CAT AAC CAA AAT ATG TCT GGA TTG 6045
 Ile Gln Ser Glu Glu Ile Leu Gln His Asn Gln Asn Met Ser Gly Leu
 1925 1930 1935

GAG AAA GTT TCT AAA ATA TCA CCT TGT GAT GTT AGT TTG GAA ACT TCA 6093
 Glu Lys Val Ser Lys Ile Ser Pro Cys Asp Val Ser Leu Glu Thr Ser
 1940 1945 1950 1955

GAT ATA TGT AAA TGT AGT ATA GGG AAG CTT CAT AAG TCA GTC TCA TCT 6141
 Asp Ile Cys Lys Cys Ser Ile Gly Lys Leu His Lys Ser Val Ser Ser
 1960 1965 1970

GCA AAT ACT TGT GGG ATT TTT AGC ACA GCA AGT GGA AAA TCT GTC CAG 6189
 Ala Asn Thr Cys Gly Ile Phe Ser Thr Ala Ser Gly Lys Ser Val Gln
 1975 1980 1985

GTA TCA GAT GCT TCA TTA CAA AAC GCA AGA CAA GTG TTT TCT GAA ATA 6237
 Val Ser Asp Ala Ser Leu Gln Asn Ala Arg Gln Val Phe Ser Glu Ile
 1990 1995 2000

GAA GAT AGT ACC AAG CAA GTC TTT TCC AAA GTA TTG TTT AAA AGT AAC 6285
 Glu Asp Ser Thr Lys Gln Val Phe Ser Lys Val Leu Phe Lys Ser Asn
 2005 2010 2015

| | |
|---|------|
| GAA CAT TCA GAC CAG CTC ACA AGA GAA GAA AAT ACT GCT ATA CGT ACT Glu His Ser Asp Gln Leu Thr Arg Glu Glu Asn Thr Ala Ile Arg Thr 2020 2025 2030 2035 | 6333 |
| CCA GAA CAT TTA ATA TCC CAA AAA GGC TTT TCA TAT AAT GTG GTA AAT Pro Glu His Leu Ile Ser Gln Lys Gly Phe Ser Tyr Asn Val Val Asn 2040 2045 2050 | 6381 |
| TCA TCT GCT TTC TCT GGA TTT AGT ACA GCA AGT GGA AAG CAA GTT TCC Ser Ser Ala Phe Ser Gly Phe Ser Thr Ala Ser Gly Lys Gln Val Ser 2055 2060 2065 | 6429 |
| ATT TTA GAA AGT TCC TTA CAC AAA GTT AAG GGA GTG TTA GAG GAA TTT Ile Leu Glu Ser Ser Leu His Lys Val Lys Gly Val Leu Glu Glu Phe 2070 2075 2080 | 6477 |
| GAT TTA ATC AGA ACT GAG CAT AGT CTT CAC TAT TCA CCT ACG TCT AGA Asp Leu Ile Arg Thr Glu His Ser Leu His Tyr Ser Pro Thr Ser Arg 2085 2090 2095 | 6525 |
| CAA AAT GTA TCA AAA ATA CTT CCT CGT GTT GAT AAG AGA AAC CCA GAG Gln Asn Val Ser Lys Ile Leu Pro Arg Val Asp Lys Arg Asn Pro Glu 2100 2105 2110 2115 | 6573 |
| CAC TGT GTA AAC TCA GAA ATG GAA AAA ACC TGC AGT AAA GAA TTT AAA His Cys Val Asn Ser Glu Met Glu Lys Thr Cys Ser Lys Glu Phe Lys 2120 2125 2130 | 6621 |
| TTA TCA AAT AAC TTA AAT GTT GAA GGT GGT TCT TCA GAA AAT AAT CAC Leu Ser Asn Asn Leu Asn Val Glu Gly Ser Ser Glu Asn Asn His 2135 2140 2145 | 6669 |
| TCT ATT AAA GTT TCT CCA TAT CTC TCT CAA TTT CAA CAA GAC AAA CAA Ser Ile Lys Val Ser Pro Tyr Leu Ser Gln Phe Gln Gln Asp Lys Gln 2150 2155 2160 | 6717 |
| CAG TTG GTA TTA GGA ACC AAA GTC TCA CTT GTT GAG AAC ATT CAT GTT Gln Leu Val Leu Gly Thr Lys Val Ser Leu Val Glu Asn Ile His Val 2165 2170 2175 | 6765 |
| TTG GGA AAA GAA CAG GCT TCA CCT AAA AAC GTA AAA ATG GAA ATT GGT Leu Gly Lys Glu Gln Ala Ser Pro Lys Asn Val Lys Met Glu Ile Gly 2180 2185 2190 2195 | 6813 |
| AAA ACT GAA ACT TTT TCT GAT GTT CCT GTG AAA ACA AAT ATA GAA GTT Lys Thr Glu Thr Phe Ser Asp Val Pro Val Lys Thr Asn Ile Glu Val 2200 2205 2210 | 6861 |
| TGT TCT ACT TAC TCC AAA GAT TCA GAA AAC TAC TTT GAA ACA GAA GCA Cys Ser Thr Tyr Ser Lys Asp Ser Glu Asn Tyr Phe Glu Thr Glu Ala 2215 2220 2225 | 6909 |
| GTA GAA ATT GCT AAA GCT TTT ATG GAA GAT GAT GAA CTG ACA GAT TCT Val Glu Ile Ala Lys Ala Phe Met Glu Asp Asp Glu Leu Thr Asp Ser 2230 2235 2240 | 6957 |
| AAA CTG CCA AGT CAT GCC ACA CAT TCT CTT TTT ACA TGT CCC GAA AAT Lys Leu Pro Ser His Ala Thr His Ser Leu Phe Thr Cys Pro Glu Asn 2245 2250 2255 | 7005 |
| GAG GAA ATG GTT TTG TCA AAT TCA AGA ATT GGA AAA AGA AGA GGA GAG | 7053 |

Glu Glu Met Val Leu Ser Asn Ser Arg Ile Gly Lys Arg Arg Gly Glu
 2260 2265 2270 2275

CCC CTT ATC TTA GTG GGA GAA CCC TCA ATC AAA AGA AAC TTA TTA AAT
 Pro Leu Ile Leu Val Gly Glu Pro Ser Ile Lys Arg Asn Leu Leu Asn
 2280 2285 2290

7101

GAA TTT GAC AGG ATA ATA GAA AAT CAA GAA AAA TCC TTA AAG GCT TCA
 Glu Phe Asp Arg Ile Ile Glu Asn Gln Glu Lys Ser Leu Lys Ala Ser
 2295 2300 2305

7149

AAA AGC ACT CCA GAT GGC ACA ATA AAA GAT CGA AGA TTG TTT ATG CAT
 Lys Ser Thr Pro Asp Gly Thr Ile Lys Asp Arg Arg Leu Phe Met His
 2310 2315 2320

7197

CAT GTT TCT TTA GAG CCG ATT ACC TGT GTA CCC TTT CGC ACA ACT AAG
 His Val Ser Leu Glu Pro Ile Thr Cys Val Pro Phe Arg Thr Thr Lys
 2325 2330 2335

7245

GAA CGT CAA GAG ATA CAG AAT CCA AAT TTT ACC GCA CCT GGT CAA GAA
 Glu Arg Gln Glu Ile Gln Asn Pro Asn Phe Thr Ala Pro Gly Gln Glu
 2340 2345 2350 2355

7293

TTT CTG TCT AAA TCT CAT TTG TAT GAA CAT CTG ACT TTG GAA AAA TCT
 Phe Leu Ser Lys Ser His Leu Tyr Glu His Leu Thr Leu Glu Lys Ser
 2360 2365 2370

7341

TCA AGC AAT TTA GCA GTT TCA GGA CAT CCA TTT TAT CAA GTT TCT GCT
 Ser Ser Asn Leu Ala Val Ser Gly His Pro Phe Tyr Gln Val Ser Ala
 2375 2380 2385

7389

ACA AGA AAT GAA AAA ATG AGA CAC TTG ATT ACT ACA GGC AGA CCA ACC
 Thr Arg Asn Glu Lys Met Arg His Leu Ile Thr Thr Gly Arg Pro Thr
 2390 2395 2400

7437

AAA GTC TTT GTT CCA CCT TTT AAA ACT AAA TCG CAT TTT CAC AGA GTT
 Lys Val Phe Val Pro Pro Phe Lys Thr Lys Ser His Phe His Arg Val
 2405 2410 2415

7485

GAA CAG TGT GTT AGG AAT ATT AAC TTG GAG GAA AAC AGA CAA AAG CAA
 Glu Gln Cys Val Arg Asn Ile Asn Leu Glu Glu Asn Arg Gln Lys Gln
 2420 2425 2430 2435

7533

AAC ATT GAT GGA CAT GGC TCT GAT GAT AGT AAA AAT AAG ATT AAT GAC
 Asn Ile Asp Gly His Gly Ser Asp Asp Ser Lys Asn Lys Ile Asn Asp
 2440 2445 2450

7581

AAT GAG ATT CAT CAG TTT AAC AAA AAC AAC TCC AAT CAA GCA GCA GCT
 Asn Glu Ile His Gln Phe Asn Lys Asn Asn Ser Asn Gln Ala Ala Ala
 2455 2460 2465

7629

GTA ACT TTC ACA AAG TGT GAA GAA CCT TTA GAT TTA ATT ACA AGT
 Val Thr Phe Thr Lys Cys Glu Glu Pro Leu Asp Leu Ile Thr Ser
 2470 2475 2480

7677

CTT CAG AAT GCC AGA GAT ATA CAG GAT ATG CGA ATT AAG AAG AAA CAA
 Leu Gln Asn Ala Arg Asp Ile Gln Asp Met Arg Ile Lys Lys Lys Gln
 2485 2490 2495

7725

AGG CAA CGC GTC TTT CCA CAG CCA GGC AGT CTG TAT CTT GCA AAA ACA
 Arg Gln Arg Val Phe Pro Gln Pro Gly Ser Leu Tyr Leu Ala Lys Thr

7773

2500 2505 2510 2515
 TCC ACT CTG CCT CGA ATC TCT CTG AAA GCA GCA GTA GGA GGC CAA GTT 7821
 Ser Thr Leu Pro Arg Ile Ser Leu Lys Ala Ala Val Gly Gly Gln Val
 2520 2525 2530

 CCC TCT GCG TGT TCT CAT AAA CAG CTG TAT ACG TAT GGC GTT TCT AAA 7869
 Pro Ser Ala Cys Ser His Lys Gln Leu Tyr Thr Tyr Gly Val Ser Lys
 2535 2540 2545

 CAT TGC ATA AAA ATT AAC AGC AAA AAT GCA GAG TCT TTT CAG TTT CAC 7917
 His Cys Ile Lys Ile Asn Ser Lys Asn Ala Glu Ser Phe Gln Phe His
 2550 2555 2560

 ACT GAA GAT TAT TTT GGT AAG GAA AGT TTA TGG ACT GGA AAA GGA ATA 7965
 Thr Glu Asp Tyr Phe Gly Lys Glu Ser Leu Trp Thr Gly Lys Gly Ile
 2565 2570 2575

 CAG TTG GCT GAT GGT GGA TGG CTC ATA CCC TCC AAT GAT GGA AAG GCT 8013
 Gln Leu Ala Asp Gly Gly Trp Leu Ile Pro Ser Asn Asp Gly Lys Ala
 2580 2585 2590 2595

 GGA AAA GAA GAA TTT TAT AGG GCT CTG TGT GAC ACT CCA GGT GTG GAT 8061
 Gly Lys Glu Glu Phe Tyr Arg Ala Leu Cys Asp Thr Pro Gly Val Asp
 2600 2605 2610

 CCA AAG CTT ATT TCT AGA ATT TGG GTT TAT AAT CAC TAT AGA TGG ATC 8109
 Pro Lys Leu Ile Ser Arg Ile Trp Val Tyr Asn His Tyr Arg Trp Ile
 2615 2620 2625

 ATA TGG AAA CTG GCA GCT ATG GAA TGT GCC TTT CCT AAG GAA TTT GCT 8157
 Ile Trp Lys Leu Ala Ala Met Glu Cys Ala Phe Pro Lys Glu Phe Ala
 2630 2635 2640

 AAT AGA TGC CTA AGC CCA GAA AGG GTG CTT CTT CAA CTA AAA TAC AGA 8205
 Asn Arg Cys Leu Ser Pro Glu Arg Val Leu Leu Gln Leu Lys Tyr Arg
 2645 2650 2655

 TAT GAT ACG GAA ATT GAT AGA AGC AGA AGA TCG GCT ATA AAA AAG ATA 8253
 Tyr Asp Thr Glu Ile Asp Arg Ser Arg Arg Ser Ala Ile Lys Lys Ile
 2660 2665 2670 2675

 ATG GAA AGG GAT GAC ACA GCT GCA AAA ACA CTT GTT CTC TGT GTT TCT 8301
 Met Glu Arg Asp Asp Thr Ala Ala Lys Thr Leu Val Leu Cys Val Ser
 2680 2685 2690

 GAC ATA ATT TCA TTG AGC GCA AAT ATA TCT GAA ACT TCT AGC AAT AAA 8349
 Asp Ile Ile Ser Leu Ser Ala Asn Ile Ser Glu Thr Ser Ser Asn Lys
 2695 2700 2705

 ACT AGT AGT GCA GAT ACC CAA AAA GTG GCC ATT ATT GAA CTT ACA GAT 8397
 Thr Ser Ser Ala Asp Thr Gln Lys Val Ala Ile Ile Glu Leu Thr Asp
 2710 2715 2720

 GGG TGG TAT GCT GTT AAG GCC CAG TTA GAT CCT CCC CTC TTA GCT GTC 8445
 Gly Trp Tyr Ala Val Lys Ala Gln Leu Asp Pro Pro Leu Leu Ala Val
 2725 2730 2735

 TTA AAG AAT GGC AGA CTG ACA GTT GGT CAG AAG ATT ATT CTT CAT GGA 8493
 Leu Lys Asn Gly Arg Leu Thr Val Gly Gln Lys Ile Ile Leu His Gly
 2740 2745 2750 2755

D
CONK

| | |
|---|------|
| GCA GAA CTG GTG GGC TCT CCT GAT GCC TGT ACA CCT CTT GAA GCC CCA Ala Glu Leu Val Gly Ser Pro Asp Ala Cys Thr Pro Leu Glu Ala Pro 2760 2765 2770 | 8541 |
| GAA TCT CTT ATG TTA AAG ATT TCT GCT AAC AGT ACT CGG CCT GCT CGC Glu Ser Leu Met Leu Lys Ile Ser Ala Asn Ser Thr Arg Pro Ala Arg 2775 2780 2785 | 8589 |
| TGG TAT ACC AAA CTT GGA TTC TTT CCT GAC CCT AGA CCT TTT CCT CTG Trp Tyr Thr Lys Leu Gly Phe Phe Pro Asp Pro Arg Pro Phe Pro Leu 2790 2795 2800 | 8637 |
| CCC TTA TCA TCG CTT TTC AGT GAT GGA GGA AAT GTT GGT TGT GTT GAT Pro Leu Ser Ser Leu Phe Ser Asp Gly Gly Asn Val Gly Cys Val Asp 2805 2810 2815 | 8685 |
| GTA ATT ATT CAA AGA GCA TAC CCT ATA CAG TGG ATG GAG AAG ACA TCA Val Ile Ile Gln Arg Ala Tyr Pro Ile Gln Trp Met Glu Lys Thr Ser 2820 2825 2830 2835 | 8733 |
| TCT GGA TTA TAC ATA TTT CGC AAT GAA AGA GAG GAA AAG GAA GCA Ser Gly Leu Tyr Ile Phe Arg Asn Glu Arg Glu Glu Lys Glu Ala 2840 2845 2850 | 8781 |
| GCA AAA TAT GTG GAG GCC CAA CAA AAG AGA CTA GAA GCC TTA TTC ACT Ala Lys Tyr Val Glu Ala Gln Gln Lys Arg Leu Glu Ala Leu Phe Thr 2855 2860 2865 | 8829 |
| AAA ATT CAG GAG GAA TTT GAA GAA CAT GAA GAA AAC ACA ACA AAA CCA Lys Ile Gln Glu Glu Phe Glu His Glu Glu Asn Thr Thr Lys Pro 2870 2875 2880 | 8877 |
| TAT TTA CCA TCA CGT GCA CTA ACA AGA CAG CAA GTT CGT GCT TTG CAA Tyr Leu Pro Ser Arg Ala Leu Thr Arg Gln Gln Val Arg Ala Leu Gln 2885 2890 2895 | 8925 |
| GAT GGT GCA GAG CTT TAT GAA GCA GTG AAG AAT GCA GCA GAC CCA GCT Asp Gly Ala Glu Leu Tyr Glu Ala Val Lys Asn Ala Ala Asp Pro Ala 2900 2905 2910 2915 | 8973 |
| TAC CTT GAG GGT TAT TTC AGT GAA GAG CAG TTA AGA GCC TTG AAT AAT Tyr Leu Glu Gly Tyr Phe Ser Glu Glu Gln Leu Arg Ala Leu Asn Asn 2920 2925 2930 | 9021 |
| CAC AGG CAA ATG TTG AAT GAT AAG AAA CAA GCT CAG ATC CAG TTG GAA His Arg Gln Met Leu Asn Asp Lys Lys Gln Ala Gln Ile Gln Leu Glu 2935 2940 2945 | 9069 |
| ATT AGG AAG GCC ATG GAA TCT GCT GAA CAA AAG GAA CAA GGT TTA TCA Ile Arg Lys Ala Met Glu Ser Ala Glu Gln Lys Glu Gln Gly Leu Ser 2950 2955 2960 | 9117 |
| AGG GAT GTC ACA ACC GTG TGG AAG TTG CGT ATT GTA AGC TAT TCA AAA Arg Asp Val Thr Thr Val Trp Lys Leu Arg Ile Val Ser Tyr Ser Lys 2965 2970 2975 | 9165 |
| AAA GAA AAA GAT TCA GTT ATA CTG AGT ATT TGG CGT CCA TCA TCA GAT Lys Glu Lys Asp Ser Val Ile Leu Ser Ile Trp Arg Pro Ser Ser Asp 2980 2985 2990 2995 | 9213 |

| | |
|---|------|
| TTA TAT TCT CTG TTA ACA GAA GGA AAG AGA TAC AGA ATT TAT CAT CTT Leu Tyr Ser Leu Leu Thr Glu Gly Lys Arg Tyr Arg Ile Tyr His Leu 3000 3005 3010 | 9261 |
| GCA ACT TCA AAA TCT AAA AGT AAA TCT GAA AGA GCT AAC ATA CAG TTA Ala Thr Ser Lys Ser Lys Ser Glu Arg Ala Asn Ile Gln Leu 3015 3020 3025 | 9309 |
| GCA GCG ACA AAA AAA ACT CAG TAT CAA CAA CTA CCG GTT TCA GAT GAA Ala Ala Thr Lys Lys Thr Gln Tyr Gln Gln Leu Pro Val Ser Asp Glu 3030 3035 3040 | 9357 |
| ATT TTA TTT CAG ATT TAC CAG CCA CGG GAG CCC CTT CAC TTC AGC AAA Ile Leu Phe Gln Ile Tyr Gln Pro Arg Glu Pro Leu His Phe Ser Lys 3045 3050 3055 | 9405 |
| TTT TTA GAT CCA GAC TTT CAG CCA TCT TGT TCT GAG GTG GAC CTA ATA Phe Leu Asp Pro Asp Phe Gln Pro Ser Cys Ser Glu Val Asp Leu Ile 3060 3065 3070 3075 | 9453 |
| GGA TTT GTC GTT TCT GTG AAA AAA ACA GGA CTT GCC CCT TTC GTC Gly Phe Val Val Ser Val Val Lys Lys Thr Gly Leu Ala Pro Phe Val 3080 3085 3090 | 9501 |
| TAT TTG TCA GAC GAA TGT TAC AAT TTA CTG GCA ATA AAG TTT TGG ATA Tyr Leu Ser Asp Glu Cys Tyr Asn Leu Ala Ile Lys Phe Trp Ile 3095 3100 3105 | 9549 |
| GAC CTT AAT GAG GAC ATT ATT AAG CCT CAT ATG TTA ATT GCT GCA AGC Asp Leu Asn Glu Asp Ile Ile Lys Pro His Met Leu Ile Ala Ala Ser 3110 3115 3120 | 9597 |
| AAC CTC CAG TGG CGA CCA GAA TCC AAA TCA GGC CTT CTT ACT TTA TTT Asn Leu Gln Trp Arg Pro Glu Ser Lys Ser Gly Leu Leu Thr Leu Phe 3125 3130 3135 | 9645 |
| GCT GGA GAT TTT TCT GTG TTT TCT GCT AGT CCA AAA GAG GGC CAC TTT Ala Gly Asp Phe Ser Val Phe Ser Ala Ser Pro Lys Glu Gly His Phe 3140 3145 3150 3155 | 9693 |
| CAA GAG ACA TTC AAC AAA ATG AAA AAT ACT GTT GAG AAT ATT GAC ATA Gln Glu Thr Phe Asn Lys Met Lys Asn Thr Val Glu Asn Ile Asp Ile 3160 3165 3170 | 9741 |
| CTT TGC AAT GAA GCA GAA AAC AAG CTT ATG CAT ATA CTG CAT GCA AAT Leu Cys Asn Glu Ala Glu Asn Lys Leu Met His Ile Leu His Ala Asn 3175 3180 3185 | 9789 |
| GAT CCC AAG TGG TCC ACC CCA ACT AAA GAC TGT ACT TCA GGG CCG TAC Asp Pro Lys Trp Ser Thr Pro Thr Lys Asp Cys Thr Ser Gly Pro Tyr 3190 3195 3200 | 9837 |
| ACT GCT CAA ATC ATT CCT GGT ACA GGA AAC AAG CTT CTG ATG TCT TCT Thr Ala Gln Ile Ile Pro Gly Thr Gly Asn Lys Leu Leu Met Ser Ser 3205 3210 3215 | 9885 |
| CCT AAT TGT GAG ATA TAT TAT CAA AGT CCT TTA TCA CTT TGT ATG GCC Pro Asn Cys Glu Ile Tyr Tyr Gln Ser Pro Leu Ser Leu Cys Met Ala 3220 3225 3230 3235 | 9933 |
| AAA AGG AAG TCT GTT TCC ACA CCT GTC TCA GCC CAG ATG ACT TCA AAG | 9981 |

Lys Arg Lys Ser Val Ser Thr Pro Val Ser Ala Gln Met Thr Ser Lys
 3240 3245 3250

TCT TGT AAA GGG GAG AAA GAG ATT GAT GAC CAA AAG AAC TGC AAA AAG 10029
 Ser Cys Lys Gly Glu Lys Glu Ile Asp Asp Gln Lys Asn Cys Lys Lys
 3255 3260 3265

AGA AGA GCC TTG GAT TTC TTG AGT AGA CTG CCT TTA CCT CCA CCT GTT 10077
 Arg Arg Ala Leu Asp Phe Leu Ser Arg Leu Pro Leu Pro Pro Val
 3270 3275 3280

AGT CCC ATT TGT ACA TTT GTT TCT CCG GCT GCA CAG AAG GCA TTT CAG 10125
 Ser Pro Ile Cys Thr Phe Val Ser Pro Ala Ala Gln Lys Ala Phe Gln
 3285 3290 3295

CCA CCA AGG AGT TGT GGC ACC AAA TAC GAA ACA CCC ATA AAG AAA AAA 10173
 Pro Pro Arg Ser Cys Gly Thr Lys Tyr Glu Thr Pro Ile Lys Lys Lys
 3300 3305 3310 3315

GAA CTG AAT TCT CCT CAG ATG ACT CCA TTT AAA AAA TTC AAT GAA ATT 10221
 Glu Leu Asn Ser Pro Gln Met Thr Pro Phe Lys Lys Phe Asn Glu Ile
 3320 3325 3330

TCT CTT TTG GAA AGT AAT TCA ATA GCT GAC GAA GAA CTT GCA TTG ATA 10269
 Ser Leu Leu Glu Ser Asn Ser Ile Ala Asp Glu Glu Leu Ala Leu Ile
 3335 3340 3345

AAT ACC CAA GCT CTT TTG TCT GGT TCA ACA GGA GAA AAA CAA TTT ATA 10317
 Asn Thr Gln Ala Leu Leu Ser Gly Ser Thr Gly Glu Lys Gln Phe Ile
 3350 3355 3360

TCT GTC AGT GAA TCC ACT AGG ACT GCT CCC ACC AGT TCA GAA GAT TAT 10365
 Ser Val Ser Glu Ser Thr Arg Thr Ala Pro Thr Ser Ser Glu Asp Tyr
 3365 3370 3375

CTC AGA CTG AAA CGA CGT TGT ACT ACA TCT CTG ATC AAA GAA CAG GAG 10413
 Leu Arg Leu Lys Arg Arg Cys Thr Thr Ser Leu Ile Lys Glu Gln Glu
 3380 3385 3390 3395

AGT TCC CAG GCC AGT ACG GAA GAA TGT GAG AAA AAT AAG CAG GAC ACA 10461
 Ser Ser Gln Ala Ser Thr Glu Glu Cys Glu Lys Asn Lys Gln Asp Thr
 3400 3405 3410

ATT ACA ACT AAA AAA TAT ATC TAA
 Ile Thr Thr Lys Lys Tyr Ile
 3415

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 3418 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: protein
- (v) FRAGMENT TYPE: internal

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Met Pro Ile Gly Ser Lys Glu Arg Pro Thr Phe Phe Glu Ile Phe Lys
 1 5 10 15
 Thr Arg Cys Asn Lys Ala Asp Leu Gly Pro Ile Ser Leu Asn Trp Phe
 20 25 30
 Glu Glu Leu Ser Ser Glu Ala Pro Pro Tyr Asn Ser Glu Pro Ala Glu
 35 40 45
 Glu Ser Glu His Lys Asn Asn Tyr Glu Pro Asn Leu Phe Lys Thr
 50 55 60
 Pro Gln Arg Lys Pro Ser Tyr Asn Gln Leu Ala Ser Thr Pro Ile Ile
 65 70 75 80
 Phe Lys Glu Gln Gly Leu Thr Leu Pro Leu Tyr Gln Ser Pro Val Lys
 85 90 95
 Glu Leu Asp Lys Phe Lys Leu Asp Leu Gly Arg Asn Val Pro Asn Ser
 100 105 110
 Arg His Lys Ser Leu Arg Thr Val Lys Thr Lys Met Asp Gln Ala Asp
 115 120 125
 Asp Val Ser Cys Pro Leu Leu Asn Ser Cys Leu Ser Glu Ser Pro Val
 130 135 140
 Val Leu Gln Cys Thr His Val Thr Pro Gln Arg Asp Lys Ser Val Val
 145 150 155 160
 Cys Gly Ser Leu Phe His Thr Pro Lys Phe Val Lys Gly Arg Gln Thr
 165 170 175
 Pro Lys His Ile Ser Glu Ser Leu Gly Ala Glu Val Asp Pro Asp Met
 180 185 190
 Ser Trp Ser Ser Ser Leu Ala Thr Pro Pro Thr Leu Ser Ser Thr Val
 195 200 205
 Leu Ile Val Arg Asn Glu Glu Ala Ser Glu Thr Val Phe Pro His Asp
 210 215 220
 Thr Thr Ala Asn Val Lys Ser Tyr Phe Ser Asn His Asp Glu Ser Leu
 225 230 235 240
 Lys Lys Asn Asp Arg Phe Ile Ala Ser Val Thr Asp Ser Glu Asn Thr
 245 250 255
 Asn Gln Arg Glu Ala Ala Ser His Gly Phe Gly Lys Thr Ser Gly Asn
 260 265 270
 Ser Phe Lys Val Asn Ser Cys Lys Asp His Ile Gly Lys Ser Met Pro
 275 280 285
 Asn Val Leu Glu Asp Glu Val Tyr Glu Thr Val Val Asp Thr Ser Glu
 290 295 300
 Glu Asp Ser Phe Ser Leu Cys Phe Ser Lys Cys Arg Thr Lys Asn Leu
 305 310 315 320
 Gln Lys Val Arg Thr Ser Lys Thr Arg Lys Lys Ile Phe His Glu Ala
 325 330 335
 Asn Ala Asp Glu Cys Glu Lys Ser Lys Asn Gln Val Lys Glu Lys Tyr
 340 345 350
 Ser Phe Val Ser Glu Val Glu Pro Asn Asp Thr Asp Pro Leu Asp Ser
 355 360 365
 Asn Val Ala His Gln Lys Pro Phe Glu Ser Gly Ser Asp Lys Ile Ser
 370 375 380
 Lys Glu Val Val Pro Ser Leu Ala Cys Glu Trp Ser Gln Leu Thr Leu
 385 390 395 400
 Ser Gly Leu Asn Gly Ala Gln Met Glu Lys Ile Pro Leu Leu His Ile
 405 410 415
 Ser Ser Cys Asp Gln Asn Ile Ser Glu Lys Asp Leu Leu Asp Thr Glu
 420 425 430
 Asn Lys Arg Lys Lys Asp Phe Leu Thr Ser Glu Asn Ser Leu Pro Arg
 435 440 445
 Ile Ser Ser Leu Pro Lys Ser Glu Lys Pro Leu Asn Glu Glu Thr Val
 450 455 460
 Val Asn Lys Arg Asp Glu Glu Gln His Leu Glu Ser His Thr Asp Cys
 465 470 475 480
 Ile Leu Ala Val Lys Gln Ala Ile Ser Gly Thr Ser Pro Val Ala Ser

DIX
CONX

| | 485 | | 490 | | 495 |
|---------------------------------|-----------------------------|---------------------|-----|--|-----|
| Ser Phe Gln Gly Ile Lys Lys Ser | Ile Phe Arg Ile Arg Glu | Ser Pro | | | |
| 500 | 505 | 510 | | | |
| Lys Glu Thr Phe Asn Ala Ser | Phe Ser Gly His Met Thr | Asp Pro Asn | | | |
| 515 | 520 | 525 | | | |
| Phe Lys Lys Glu Thr Glu Ala | Ser Glu Ser Gly Leu | Glu Ile His Thr | | | |
| 530 | 535 | 540 | | | |
| Val Cys Ser Gln Lys Glu Asp | Ser Leu Cys Pro Asn | Leu Ile Asp Asn | | | |
| 545 | 550 | 555 | | | |
| Gly Ser Trp Pro Ala Thr Thr | Gln Asn Ser Val Ala | Leu Lys Asn | | | |
| 565 | 570 | 575 | | | |
| Ala Gly Leu Ile Ser Thr Leu | Lys Lys Thr Asn Lys Phe | Ile Tyr | | | |
| 580 | 585 | 590 | | | |
| Ala Ile His Asp Glu Thr Ser | Tyr Lys Gly Lys Lys Ile | Pro Lys Asp | | | |
| 595 | 600 | 605 | | | |
| Gln Lys Ser Glu Leu Ile Asn | Cys Ser Ala Gln Phe | Glu Ala Asn Ala | | | |
| 610 | 615 | 620 | | | |
| Phe Glu Ala Pro Leu Thr Phe | Ala Asn Ala Asp Ser | Gly Leu Leu His | | | |
| 625 | 630 | 635 | | | |
| Ser Ser Val Lys Arg Ser Cys | Ser Gln Asn Asp Ser | Glu Glu Pro Thr | | | |
| 645 | 650 | 655 | | | |
| Leu Ser Leu Thr Ser Ser Phe | Gly Thr Ile Leu Arg Lys Cys | Ser Arg | | | |
| 660 | 665 | 670 | | | |
| Asn Glu Thr Cys Ser Asn Asn | Thr Val Ile Ser Gln Asp | Leu Asp Tyr | | | |
| 675 | 680 | 685 | | | |
| Lys Glu Ala Lys Cys Asn | Lys Glu Lys Leu Gln Leu | Phe Ile Thr Pro | | | |
| 690 | 695 | 700 | | | |
| Glu Ala Asp Ser Leu Ser Cys | Leu Gln Glu Gly Gln Cys | Glu Asn Asp | | | |
| 705 | 710 | 715 | | | |
| Pro Lys Ser Lys Lys Val Ser | Asp Ile Lys Glu Glu Val | Leu Ala Ala | | | |
| 725 | 730 | 735 | | | |
| Ala Cys His Pro Val Gln His | Ser Lys Val Glu Tyr Ser | Asp Thr Asp | | | |
| 740 | 745 | 750 | | | |
| Phe Gln Ser Gln Lys Ser Leu | Leu Tyr Asp His Glu Asn | Ala Ser Thr | | | |
| 755 | 760 | 765 | | | |
| Leu Ile Leu Thr Pro Thr Ser | Lys Asp Val Leu Ser Asn | Leu Val Met | | | |
| 770 | 775 | 780 | | | |
| Ile Ser Arg Gly Lys Glu Ser | Tyr Lys Met Ser Asp Lys Leu | Lys Gly | | | |
| 785 | 790 | 795 | | | |
| Asn Asn Tyr Glu Ser Asp Val | Glu Leu Thr Lys Asn Ile | Pro Met Glu | | | |
| 805 | 810 | 815 | | | |
| Lys Asn Gln Asp Val Cys | Ala Leu Asn Glu Asn | Tyr Lys Asn Val Glu | | | |
| 820 | 825 | 830 | | | |
| Leu Leu Pro Pro Glu Lys Tyr | Met Arg Val Ala Ser | Pro Ser Arg Lys | | | |
| 835 | 840 | 845 | | | |
| Val Gln Phe Asn Gln Asn | Thr Asn Leu Arg Val Ile | Gln Lys Asn Gln | | | |
| 850 | 855 | 860 | | | |
| Glu Glu Thr Thr Ser Ile Ser | Lys Ile Thr Val Asn Pro | Asp Ser Glu | | | |
| 865 | 870 | 875 | | | |
| Glu Leu Phe Ser Asp Asn Glu | Asn Asn Phe Val Phe | Gln Val Ala Asn | | | |
| 885 | 890 | 895 | | | |
| Glu Arg Asn Asn Leu Ala | Leu Gly Asn Thr Lys Glu | Leu His Glu Thr | | | |
| 900 | 905 | 910 | | | |
| Asp Leu Thr Cys Val Asn Glu | Pro Ile Phe Lys Asn Ser | Thr Met Val | | | |
| 915 | 920 | 925 | | | |
| Leu Tyr Gly Asp Thr Gly | Asp Lys Gln Ala Thr | Gln Val Ser Ile Lys | | | |
| 930 | 935 | 940 | | | |
| Lys Asp Leu Val Tyr Val | Leu Ala Glu Glu Asn | Lys Asn Ser Val Lys | | | |
| 945 | 950 | 955 | | | |
| Gln His Ile Lys Met Thr | Leu Gly Gln Asp Leu | Lys Ser Asp Ile Ser | | | |
| 965 | 970 | 975 | | | |

Leu Asn Ile Asp Lys Ile Pro Glu Lys Asn Asn Asp Tyr Met Asn Lys
 980 985 990
 Trp Ala Gly Leu Leu Gly Pro Ile Ser Asn His Ser Phe Gly Gly Ser
 995 1000 1005
 Phe Arg Thr Ala Ser Asn Lys Glu Ile Lys Leu Ser Glu His Asn Ile
 1010 1015 1020
 Lys Lys Ser Lys Met Phe Phe Lys Asp Ile Glu Glu Gln Tyr Pro Thr
 1025 1030 1035 1040
 Ser Leu Ala Cys Val Glu Ile Val Asn Thr Leu Ala Leu Asp Asn Gln
 1045 1050 1055
 Lys Lys Leu Ser Lys Pro Gln Ser Ile Asn Thr Val Ser Ala His Leu
 1060 1065 1070
 Gln Ser Ser Val Val Val Ser Asp Cys Lys Asn Ser His Ile Thr Pro
 1075 1080 1085
 Gln Met Leu Phe Ser Lys Gln Asp Phe Asn Ser Asn His Asn Leu Thr
 1090 1095 1100
 Pro Ser Gln Lys Ala Glu Ile Thr Glu Leu Ser Thr Ile Leu Glu Glu
 1105 1110 1115 1120
 Ser Gly Ser Gln Phe Glu Phe Thr Gln Phe Arg Lys Pro Ser Tyr Ile
 1125 1130 1135
 Leu Gln Lys Ser Thr Phe Glu Val Pro Glu Asn Gln Met Thr Ile Leu
 1140 1145 1150
 Lys Thr Thr Ser Glu Glu Cys Arg Asp Ala Asp Leu His Val Ile Met
 1155 1160 1165
 Asn Ala Pro Ser Ile Gly Gln Val Asp Ser Ser Lys Gln Phe Glu Gly
 1170 1175 1180
 Thr Val Glu Ile Lys Arg Lys Phe Ala Gly Leu Leu Lys Asn Asp Cys
 1185 1190 1195 1200
 Asn Lys Ser Ala Ser Gly Tyr Leu Thr Asp Glu Asn Glu Val Gly Phe
 1205 1210 1215
 Arg Gly Phe Tyr Ser Ala His Gly Thr Lys Leu Asn Val Ser Thr Glu
 1220 1225 1230
 Ala Leu Gln Lys Ala Val Lys Leu Phe Ser Asp Ile Glu Asn Ile Ser
 1235 1240 1245
 Glu Glu Thr Ser Ala Glu Val His Pro Ile Ser Leu Ser Ser Ser Lys
 1250 1255 1260
 Cys His Asp Ser Val Val Ser Met Phe Lys Ile Glu Asn His Asn Asp
 1265 1270 1275 1280
 Lys Thr Val Ser Glu Lys Asn Asn Lys Cys Gln Leu Ile Leu Gln Asn
 1285 1290 1295
 Asn Ile Glu Met Thr Thr Gly Thr Phe Val Glu Glu Ile Thr Glu Asn
 1300 1305 1310
 Tyr Lys Arg Asn Thr Glu Asn Glu Asp Asn Lys Tyr Thr Ala Ala Ser
 1315 1320 1325
 Arg Asn Ser His Asn Leu Glu Phe Asp Gly Ser Asp Ser Ser Lys Asn
 1330 1335 1340
 Asp Thr Val Cys Ile His Lys Asp Glu Thr Asp Leu Leu Phe Thr Asp
 1345 1350 1355 1360
 Gln His Asn Ile Cys Leu Lys Leu Ser Gly Gln Phe Met Lys Glu Gly
 1365 1370 1375
 Asn Thr Gln Ile Lys Glu Asp Leu Ser Asp Leu Thr Phe Leu Glu Val
 1380 1385 1390
 Ala Lys Ala Gln Glu Ala Cys His Gly Asn Thr Ser Asn Lys Glu Gln
 1395 1400 1405
 Leu Thr Ala Thr Lys Thr Glu Gln Asn Ile Lys Asp Phe Glu Thr Ser
 1410 1415 1420
 Asp Thr Phe Phe Gln Thr Ala Ser Gly Lys Asn Ile Ser Val Ala Lys
 1425 1430 1435 1440
 Glu Ser Phe Asn Lys Ile Val Asn Phe Phe Asp Gln Lys Pro Glu Glu
 1445 1450 1455
 Leu His Asn Phe Ser Leu Asn Ser Glu Leu His Ser Asp Ile Arg Lys

D
cont

1460 1465 1470
 Asn Lys Met Asp Ile Leu Ser Tyr Glu Glu Thr Asp Ile Val Lys His
 1475 1480 1485
 Lys Ile Leu Lys Glu Ser Val Pro Val Gly Thr Gly Asn Gln Leu Val
 1490 1495 1500
 Thr Phe Gln Gly Gln Pro Glu Arg Asp Glu Lys Ile Lys Glu Pro Thr
 1505 1510 1515 1520
 Leu Leu Gly Phe His Thr Ala Ser Gly Lys Lys Val Lys Ile Ala Lys
 1525 1530 1535
 Glu Ser Leu Asp Lys Val Lys Asn Leu Phe Asp Glu Lys Glu Gln Gly
 1540 1545 1550
 Thr Ser Glu Ile Thr Ser Phe Ser His Gln Trp Ala Lys Thr Leu Lys
 1555 1560 1565
 Tyr Arg Glu Ala Cys Lys Asp Leu Glu Leu Ala Cys Glu Thr Ile Glu
 1570 1575 1580
 Ile Thr Ala Ala Pro Lys Cys Lys Glu Met Gln Asn Ser Leu Asn Asn
 1585 1590 1595 1600
 Asp Lys Asn Leu Val Ser Ile Glu Thr Val Val Pro Pro Lys Leu Leu
 1605 1610 1615
 Ser Asp Asn Leu Cys Arg Gln Thr Glu Asn Leu Lys Thr Ser Lys Ser
 1620 1625 1630
 Ile Phe Leu Lys Val Lys Val His Glu Asn Val Glu Lys Glu Thr Ala
 1635 1640 1645
 Lys Ser Pro Ala Thr Cys Tyr Thr Asn Gln Ser Pro Tyr Ser Val Ile
 1650 1655 1660
 Glu Asn Ser Ala Leu Ala Phe Tyr Thr Ser Cys Ser Arg Lys Thr Ser
 1665 1670 1675 1680
 Val Ser Gln Thr Ser Leu Leu Glu Ala Lys Lys Trp Leu Arg Glu Gly
 1685 1690 1695
 Ile Phe Asp Gly Gln Pro Glu Arg Ile Asn Thr Ala Asp Tyr Val Gly
 1700 1705 1710
 Asn Tyr Leu Tyr Glu Asn Asn Ser Asn Ser Thr Ile Ala Glu Asn Asp
 1715 1720 1725
 Lys Asn His Leu Ser Glu Lys Gln Asp Thr Tyr Leu Ser Asn Ser Ser
 1730 1735 1740
 Met Ser Asn Ser Tyr Ser Tyr His Ser Asp Glu Val Tyr Asn Asp Ser
 1745 1750 1755 1760
 Gly Tyr Leu Ser Lys Asn Lys Leu Asp Ser Gly Ile Glu Pro Val Leu
 1765 1770 1775
 Lys Asn Val Glu Asp Gln Lys Asn Thr Ser Phe Ser Lys Val Ile Ser
 1780 1785 1790
 Asn Val Lys Asp Ala Asn Ala Tyr Pro Gln Thr Val Asn Glu Asp Ile
 1795 1800 1805
 Cys Val Glu Leu Val Thr Ser Ser Pro Cys Lys Asn Lys Asn
 1810 1815 1820
 Ala Ala Ile Lys Leu Ser Ile Ser Asn Ser Asn Asn Phe Glu Val Gly
 1825 1830 1835 1840
 Pro Pro Ala Phe Arg Ile Ala Ser Gly Lys Ile Val Cys Val Ser His
 1845 1850 1855
 Glu Thr Ile Lys Lys Val Lys Asp Ile Phe Thr Asp Ser Phe Ser Lys
 1860 1865 1870
 Val Ile Lys Glu Asn Asn Glu Asn Lys Ser Lys Ile Cys Gln Thr Lys
 1875 1880 1885
 Ile Met Ala Gly Cys Tyr Glu Ala Leu Asp Asp Ser Glu Asp Ile Leu
 1890 1895 1900
 His Asn Ser Leu Asp Asn Asp Glu Cys Ser Thr His Ser His Lys Val
 1905 1910 1915 1920
 Phe Ala Asp Ile Gln Ser Glu Glu Ile Leu Gln His Asn Gln Asn Met
 1925 1930 1935
 Ser Gly Leu Glu Lys Val Ser Lys Ile Ser Pro Cys Asp Val Ser Leu
 1940 1945 1950

Glu Thr Ser Asp Ile Cys Lys Cys Ser Ile Gly Lys Leu His Lys Ser
 1955 1960 1965
 Val Ser Ser Ala Asn Thr Cys Gly Ile Phe Ser Thr Ala Ser Gly Lys
 1970 1975 1980
 Ser Val Gln Val Ser Asp Ala Ser Leu Gln Asn Ala Arg Gln Val Phe
 1985 1990 1995 2000
 Ser Glu Ile Glu Asp Ser Thr Lys Gln Val Phe Ser Lys Val Leu Phe
 2005 2010 2015
 Lys Ser Asn Glu His Ser Asp Gln Leu Thr Arg Glu Glu Asn Thr Ala
 2020 2025 2030
 Ile Arg Thr Pro Glu His Leu Ile Ser Gln Lys Gly Phe Ser Tyr Asn
 2035 2040 2045
 Val Val Asn Ser Ser Ala Phe Ser Gly Phe Ser Thr Ala Ser Gly Lys
 2050 2055 2060
 Gln Val Ser Ile Leu Glu Ser Ser Leu His Lys Val Lys Gly Val Leu
 2065 2070 2075 2080
 Glu Glu Phe Asp Leu Ile Arg Thr Glu His Ser Leu His Tyr Ser Pro
 2085 2090 2095
 Thr Ser Arg Gln Asn Val Ser Lys Ile Leu Pro Arg Val Asp Lys Arg
 2100 2105 2110
 Asn Pro Glu His Cys Val Asn Ser Glu Met Glu Lys Thr Cys Ser Lys
 2115 2120 2125
 Glu Phe Lys Leu Ser Asn Asn Leu Asn Val Glu Gly Gly Ser Ser Glu
 2130 2135 2140
 Asn Asn His Ser Ile Lys Val Ser Pro Tyr Leu Ser Gln Phe Gln Gln
 2145 2150 2155 2160
 Asp Lys Gln Gln Leu Val Leu Gly Thr Lys Val Ser Leu Val Glu Asn
 2165 2170 2175
 Ile His Val Leu Gly Lys Glu Gln Ala Ser Pro Lys Asn Val Lys Met
 2180 2185 2190
 Glu Ile Gly Lys Thr Glu Thr Phe Ser Asp Val Pro Val Lys Thr Asn
 2195 2200 2205
 Ile Glu Val Cys Ser Thr Tyr Ser Lys Asp Ser Glu Asn Tyr Phe Glu
 2210 2215 2220
 Thr Glu Ala Val Glu Ile Ala Lys Ala Phe Met Glu Asp Asp Glu Leu
 2225 2230 2235 2240
 Thr Asp Ser Lys Leu Pro Ser His Ala Thr His Ser Leu Phe Thr Cys
 2245 2250 2255
 Pro Glu Asn Glu Glu Met Val Leu Ser Asn Ser Arg Ile Gly Lys Arg
 2260 2265 2270
 Arg Gly Glu Pro Leu Ile Leu Val Gly Glu Pro Ser Ile Lys Arg Asn
 2275 2280 2285
 Leu Leu Asn Glu Phe Asp Arg Ile Ile Glu Asn Gln Glu Lys Ser Leu
 2290 2295 2300
 Lys Ala Ser Lys Ser Thr Pro Asp Gly Thr Ile Lys Asp Arg Arg Leu
 2305 2310 2315 2320
 Phe Met His His Val Ser Leu Glu Pro Ile Thr Cys Val Pro Phe Arg
 2325 2330 2335
 Thr Thr Lys Glu Arg Gln Glu Ile Gln Asn Pro Asn Phe Thr Ala Pro
 2340 2345 2350
 Gly Gln Glu Phe Leu Ser Lys Ser His Leu Tyr Glu His Leu Thr Leu
 2355 2360 2365
 Glu Lys Ser Ser Ser Asn Leu Ala Val Ser Gly His Pro Phe Tyr Gln
 2370 2375 2380
 Val Ser Ala Thr Arg Asn Glu Lys Met Arg His Leu Ile Thr Thr Gly
 2385 2390 2395 2400
 Arg Pro Thr Lys Val Phe Val Pro Pro Phe Lys Thr Lys Ser His Phe
 2405 2410 2415
 His Arg Val Glu Gln Cys Val Arg Asn Ile Asn Leu Glu Glu Asn Arg
 2420 2425 2430
 Gln Lys Gln Asn Ile Asp Gly His Gly Ser Asp Asp Ser Lys Asn Lys

D
K
cont

2435 2440 2445
 Ile Asn Asp Asn Glu Ile His Gln Phe Asn Lys Asn Asn Ser Asn Gln
 2450 2455 2460
 Ala Ala Ala Val Thr Phe Thr Lys Cys Glu Glu Pro Leu Asp Leu
 2465 2470 2475 2480
 Ile Thr Ser Leu Gln Asn Ala Arg Asp Ile Gln Asp Met Arg Ile Lys
 2485 2490 2495
 Lys Lys Gln Arg Gln Arg Val Phe Pro Gln Pro Gly Ser Leu Tyr Leu
 2500 2505 2510
 Ala Lys Thr Ser Thr Leu Pro Arg Ile Ser Leu Lys Ala Ala Val Gly
 2515 2520 2525
 Gly Gln Val Pro Ser Ala Cys Ser His Lys Gln Leu Tyr Thr Tyr Gly
 2530 2535 2540
 Val Ser Lys His Cys Ile Lys Ile Asn Ser Lys Asn Ala Glu Ser Phe
 2545 2550 2555 2560
 Gln Phe His Thr Glu Asp Tyr Phe Gly Lys Glu Ser Leu Trp Thr Gly
 2565 2570 2575
 Lys Gly Ile Gln Leu Ala Asp Gly Gly Trp Leu Ile Pro Ser Asn Asp
 2580 2585 2590
 Gly Lys Ala Gly Lys Glu Glu Phe Tyr Arg Ala Leu Cys Asp Thr Pro
 2595 2600 2605
 Gly Val Asp Pro Lys Leu Ile Ser Arg Ile Trp Val Tyr Asn His Tyr
 2610 2615 2620
 Arg Trp Ile Ile Trp Lys Leu Ala Ala Met Glu Cys Ala Phe Pro Lys
 2625 2630 2635 2640
 Glu Phe Ala Asn Arg Cys Leu Ser Pro Glu Arg Val Leu Leu Gln Leu
 2645 2650 2655
 Lys Tyr Arg Tyr Asp Thr Glu Ile Asp Arg Ser Arg Arg Ser Ala Ile
 2660 2665 2670
 Lys Lys Ile Met Glu Arg Asp Asp Thr Ala Ala Lys Thr Leu Val Leu
 2675 2680 2685
 Cys Val Ser Asp Ile Ile Ser Leu Ser Ala Asn Ile Ser Glu Thr Ser
 2690 2695 2700
 Ser Asn Lys Thr Ser Ser Ala Asp Thr Gln Lys Val Ala Ile Ile Glu
 2705 2710 2715 2720
 Leu Thr Asp Gly Trp Tyr Ala Val Lys Ala Gln Leu Asp Pro Pro Leu
 2725 2730 2735
 Leu Ala Val Leu Lys Asn Gly Arg Leu Thr Val Gly Gln Lys Ile Ile
 2740 2745 2750
 Leu His Gly Ala Glu Leu Val Gly Ser Pro Asp Ala Cys Thr Pro Leu
 2755 2760 2765
 Glu Ala Pro Glu Ser Leu Met Leu Lys Ile Ser Ala Asn Ser Thr Arg
 2770 2775 2780
 Pro Ala Arg Trp Tyr Thr Lys Leu Gly Phe Phe Pro Asp Pro Arg Pro
 2785 2790 2795 2800
 Phe Pro Leu Pro Leu Ser Ser Leu Phe Ser Asp Gly Gly Asn Val Gly
 2805 2810 2815
 Cys Val Asp Val Ile Ile Gln Arg Ala Tyr Pro Ile Gln Trp Met Glu
 2820 2825 2830
 Lys Thr Ser Ser Gly Leu Tyr Ile Phe Arg Asn Glu Arg Glu Glu Glu
 2835 2840 2845
 Lys Glu Ala Ala Lys Tyr Val Glu Ala Gln Gln Lys Arg Leu Glu Ala
 2850 2855 2860
 Leu Phe Thr Lys Ile Gln Glu Glu Phe Glu Glu His Glu Glu Asn Thr
 2865 2870 2875 2880
 Thr Lys Pro Tyr Leu Pro Ser Arg Ala Leu Thr Arg Gln Gln Val Arg
 2885 2890 2895
 Ala Leu Gln Asp Gly Ala Glu Leu Tyr Glu Ala Val Lys Asn Ala Ala
 2900 2905 2910
 Asp Pro Ala Tyr Leu Glu Gly Tyr Phe Ser Glu Glu Gln Leu Arg Ala
 2915 2920 2925

Leu Asn Asn His Arg Gln Met Leu Asn Asp Lys Lys Gln Ala Gln Ile
 2930 2935 2940
 Gln Leu Glu Ile Arg Lys Ala Met Glu Ser Ala Glu Gln Lys Glu Gln
 2945 2950 2955 2960
 Gly Leu Ser Arg Asp Val Thr Thr Val Trp Lys Leu Arg Ile Val Ser
 2965 2970 2975
 Tyr Ser Lys Lys Glu Lys Asp Ser Val Ile Leu Ser Ile Trp Arg Pro
 2980 2985 2990
 Ser Ser Asp Leu Tyr Ser Leu Leu Thr Glu Gly Lys Arg Tyr Arg Ile
 2995 3000 3005
 Tyr His Leu Ala Thr Ser Lys Ser Lys Ser Glu Arg Ala Asn
 3010 3015 3020
 Ile Gln Leu Ala Ala Thr Lys Lys Thr Gln Tyr Gln Gln Leu Pro Val
 3025 3030 3035 3040
 Ser Asp Glu Ile Leu Phe Gln Ile Tyr Gln Pro Arg Glu Pro Leu His
 3045 3050 3055
 Phe Ser Lys Phe Leu Asp Pro Asp Phe Gln Pro Ser Cys Ser Glu Val
 3060 3065 3070
 Asp Leu Ile Gly Phe Val Val Ser Val Val Lys Lys Thr Gly Leu Ala
 3075 3080 3085
 Pro Phe Val Tyr Leu Ser Asp Glu Cys Tyr Asn Leu Leu Ala Ile Lys
 3090 3095 3100
 Phe Trp Ile Asp Leu Asn Glu Asp Ile Ile Lys Pro His Met Leu Ile
 3105 3110 3115 3120
 Ala Ala Ser Asn Leu Gln Trp Arg Pro Glu Ser Lys Ser Gly Leu Leu
 3125 3130 3135
 Thr Leu Phe Ala Gly Asp Phe Ser Val Phe Ser Ala Ser Pro Lys Glu
 3140 3145 3150
 Gly His Phe Gln Glu Thr Phe Asn Lys Met Lys Asn Thr Val Glu Asn
 3155 3160 3165
 Ile Asp Ile Leu Cys Asn Glu Ala Glu Asn Lys Leu Met His Ile Leu
 3170 3175 3180
 His Ala Asn Asp Pro Lys Trp Ser Thr Pro Thr Lys Asp Cys Thr Ser
 3185 3190 3195 3200
 Gly Pro Tyr Thr Ala Gln Ile Ile Pro Gly Thr Gly Asn Lys Leu Leu
 3205 3210 3215
 Met Ser Ser Pro Asn Cys Glu Ile Tyr Tyr Gln Ser Pro Leu Ser Leu
 3220 3225 3230
 Cys Met Ala Lys Arg Lys Ser Val Ser Thr Pro Val Ser Ala Gln Met
 3235 3240 3245
 Thr Ser Lys Ser Cys Lys Gly Glu Lys Glu Ile Asp Asp Gln Lys Asn
 3250 3255 3260
 Cys Lys Lys Arg Arg Ala Leu Asp Phe Leu Ser Arg Leu Pro Leu Pro
 3265 3270 3275 3280
 Pro Pro Val Ser Pro Ile Cys Thr Phe Val Ser Pro Ala Ala Gln Lys
 3285 3290 3295
 Ala Phe Gln Pro Pro Arg Ser Cys Gly Thr Lys Tyr Glu Thr Pro Ile
 3300 3305 3310
 Lys Lys Lys Glu Leu Asn Ser Pro Gln Met Thr Pro Phe Lys Lys Phe
 3315 3320 3325
 Asn Glu Ile Ser Leu Leu Glu Ser Asn Ser Ile Ala Asp Glu Glu Leu
 3330 3335 3340
 Ala Leu Ile Asn Thr Gln Ala Leu Leu Ser Gly Ser Thr Gly Glu Lys
 3345 3350 3355 3360
 Gln Phe Ile Ser Val Ser Glu Ser Thr Arg Thr Ala Pro Thr Ser Ser
 3365 3370 3375
 Glu Asp Tyr Leu Arg Leu Lys Arg Arg Cys Thr Thr Ser Leu Ile Lys
 3380 3385 3390
 Glu Gln Glu Ser Ser Gln Ala Ser Thr Glu Glu Cys Glu Lys Asn Lys
 3395 3400 3405
 Gln Asp Thr Ile Thr Thr Lys Tyr Ile

D
Cont

3410

3415

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 10485 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: Coding Sequence
- (B) LOCATION: 229...10482
- (D) OTHER INFORMATION: BRCA2 (OMI5)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

D
CK
CMT

| | |
|--|-------------|
| GGTGGCGCGA GCTTCTGAAA CTAGGCGGCA GAGGC GGAGC CGCTGTGGCA CTGCTGCGCC | 60 |
| TCTGCTGCGC CTCGGGTGTC TTTTGC GGCG GTGGGTGCGC GCCGGGAGAA GCGTGAGGGG | 120 |
| ACAGATTGT GACCGGCGCG GTTTTGTCA GCTTACTCCG GCCAAAAAAAG AACTGCACCT | 180 |
| CTGGAGCGGA CTTATTTACC AAGCATTGGA GGAATATCGT AGGTAAAA ATG CCT ATT | 237 |
| | Met Pro Ile |
| | 1 |
| GGA TCC AAA GAG AGG CCA ACA TTT TTT GAA ATT TTT AAG ACA CGC TGC | 285 |
| Gly Ser Lys Glu Arg Pro Thr Phe Phe Glu Ile Phe Lys Thr Arg Cys | |
| 5 10 15 | |
| AAC AAA GCA GAT TTA GGA CCA ATA AGT CTT AAT TGG TTT GAA GAA CTT | 333 |
| Asn Lys Ala Asp Leu Gly Pro Ile Ser Leu Asn Trp Phe Glu Glu Leu | |
| 20 25 30 35 | |
| TCT TCA GAA GCT CCA CCC TAT AAT TCT GAA CCT GCA GAA GAA TCT GAA | 381 |
| Ser Ser Glu Ala Pro Pro Tyr Asn Ser Glu Pro Ala Glu Glu Ser Glu | |
| 40 45 50 | |
| CAT AAA AAC AAC AAT TAC GAA CCA AAC CTA TTT AAA ACT CCA CAA AGG | 429 |
| His Lys Asn Asn Asn Tyr Glu Pro Asn Leu Phe Lys Thr Pro Gln Arg | |
| 55 60 65 | |
| AAA CCA TCT TAT AAT CAG CTG GCT TCA ACT CCA ATA ATA TTC AAA GAG | 477 |
| Lys Pro Ser Tyr Asn Gln Leu Ala Ser Thr Pro Ile Ile Phe Lys Glu | |
| 70 75 80 | |
| CAA GGG CTG ACT CTG CCG CTG TAC CAA TCT CCT GTA AAA GAA TTA GAT | 525 |
| Gln Gly Leu Thr Leu Pro Leu Tyr Gln Ser Pro Val Lys Glu Leu Asp | |
| 85 90 95 | |
| AAA TTC AAA TTA GAC TTA GGA AGG AAT GTT CCC AAT AGT AGA CAT AAA | 573 |
| Lys Phe Lys Leu Asp Leu Gly Arg Asn Val Pro Asn Ser Arg His Lys | |
| 100 105 110 115 | |
| AGT CTT CGC ACA GTG AAA ACT AAA ATG GAT CAA GCA GAT GAT GTT TCC | 621 |
| Ser Leu Arg Thr Val Lys Thr Lys Met Asp Gln Ala Asp Asp Val Ser | |
| 120 125 130 | |
| TGT CCA CTT CTA AAT TCT TGT CTT AGT GAA AGT CCT GTT GTT CTA CAA | 669 |
| Cys Pro Leu Leu Asn Ser Cys Leu Ser Glu Ser Pro Val Val Leu Gln | |

| 135 | 140 | 145 | |
|---|-----|-----|------|
| TGT ACA CAT GTA ACA CCA CAA AGA GAT AAG TCA GTG GTA TGT GGG AGT Cys Thr His Val Thr Pro Gln Arg Asp Lys Ser Val Val Cys Gly Ser 150 | 155 | 160 | 717 |
| TTG TTT CAT ACA CCA AAG TTT GTG AAG GGT CGT CAG ACA CCA AAA CAT Leu Phe His Thr Pro Lys Phe Val Lys Gly Arg Gln Thr Pro Lys His 165 | 170 | 175 | 765 |
| ATT TCT GAA AGT CTA GGA GCT GAG GTG GAT CCT GAT ATG TCT TGG TCA Ile Ser Glu Ser Leu Gly Ala Glu Val Asp Pro Asp Met Ser Trp Ser 180 | 185 | 190 | 813 |
| AGT TCT TTA GCT ACA CCA CCC ACC CTT AGT TCT ACT GTG CTC ATA GTC Ser Ser Leu Ala Thr Pro Pro Thr Leu Ser Ser Thr Val Leu Ile Val 200 | 205 | 210 | 861 |
| AGA AAT GAA GAA GCA TCT GAA ACT GTA TTT CCT CAT GAT ACT ACT GCT Arg Asn Glu Glu Ala Ser Glu Thr Val Phe Pro His Asp Thr Thr Ala 215 | 220 | 225 | 909 |
| AAT GTG AAA AGC TAT TTT TCC AAT CAT GAT GAA AGT CTG AAG AAA AAT Asn Val Lys Ser Tyr Phe Ser Asn His Asp Glu Ser Leu Lys Lys Asn 230 | 235 | 240 | 957 |
| GAT AGA TTT ATC GCT TCT GTG ACA GAC AGT GAA AAC ACA AAT CAA AGA Asp Arg Phe Ile Ala Ser Val Thr Asp Ser Glu Asn Thr Asn Gln Arg 245 | 250 | 255 | 1005 |
| GAA GCT GCA AGT CAT GGA TTT GGA AAA ACA TCA GGG AAT TCA TTT AAA Glu Ala Ala Ser His Gly Phe Gly Lys Thr Ser Gly Asn Ser Phe Lys 260 | 265 | 270 | 1053 |
| GTA AAT AGC TGC AAA GAC CAC ATT GGA AAG TCA ATG CCA CAT GTC CTA Val Asn Ser Cys Lys Asp His Ile Gly Lys Ser Met Pro His Val Leu 280 | 285 | 290 | 1101 |
| GAA GAT GAA GTA TAT GAA ACA GTT GTA GAT ACC TCT GAA GAA GAT AGT Glu Asp Glu Val Tyr Glu Thr Val Val Asp Thr Ser Glu Glu Asp Ser 295 | 300 | 305 | 1149 |
| TTT TCA TTA TGT TTT TCT AAA TGT AGA ACA AAA AAT CTA CAA AAA GTA Phe Ser Leu Cys Phe Ser Lys Cys Arg Thr Lys Asn Leu Gln Lys Val 310 | 315 | 320 | 1197 |
| AGA ACT AGC AAG ACT AGG AAA AAA ATT TTC CAT GAA GCA AAC GCT GAT Arg Thr Ser Lys Thr Arg Lys Lys Ile Phe His Glu Ala Asn Ala Asp 325 | 330 | 335 | 1245 |
| GAA TGT GAA AAA TCT AAA AAC CAA GTG AAA GAA AAA TAC TCA TTT GTA Glu Cys Glu Lys Ser Lys Asn Gln Val Lys Glu Lys Tyr Ser Phe Val 340 | 345 | 350 | 1293 |
| TCT GAA GTG GAA CCA AAT GAT ACT GAT CCA TTA GAT TCA AAT GTA GCA Ser Glu Val Glu Pro Asn Asp Thr Asp Pro Leu Asp Ser Asn Val Ala 360 | 365 | 370 | 1341 |
| CAT CAG AAG CCC TTT GAG AGT GGA AGT GAC AAA ATC TCC AAG GAA GTT His Gln Lys Pro Phe Glu Ser Gly Ser Asp Lys Ile Ser Lys Glu Val 375 | 380 | 385 | 1389 |

| | |
|---|------|
| GTA CCG TCT TTG GCC TGT GAA TGG TCT CAA CTA ACC CTT TCA GGT CTA Val Pro Ser Leu Ala Cys Glu Trp Ser Gln Leu Thr Leu Ser Gly Leu 390 395 400 | 1437 |
| AAT GGA GCC CAG ATG GAG AAA ATA CCC CTA TTG CAT ATT TCT TCA TGT Asn Gly Ala Gln Met Glu Lys Ile Pro Leu Leu His Ile Ser Ser Cys 405 410 415 | 1485 |
| GAC CAA AAT ATT TCA GAA AAA GAC CTA TTA GAC ACA GAG AAC AAA AGA Asp Gln Asn Ile Ser Glu Lys Asp Leu Leu Asp Thr Glu Asn Lys Arg 420 425 430 435 | 1533 |
| AAG AAA GAT TTT CTT ACT TCA GAG AAT TCT TTG CCA CGT ATT TCT AGC Lys Lys Asp Phe Leu Thr Ser Glu Asn Ser Leu Pro Arg Ile Ser Ser 440 445 450 | 1581 |
| CTA CCA AAA TCG GAG AAG CCA TTA AAT GAG GAA ACA GTG GTA AAT AAG Leu Pro Lys Ser Glu Lys Pro Leu Asn Glu Glu Thr Val Val Asn Lys 455 460 465 | 1629 |
| AGA GAT GAA GAG CAG CAT CTT GAA TCT CAT ACA GAC TGC ATT CTT GCA Arg Asp Glu Glu Gln His Leu Glu Ser His Thr Asp Cys Ile Leu Ala 470 475 480 | 1677 |
| GTA AAG CAG GCA ATA TCT GGA ACT TCT CCA GTG GCT TCT TCA TTT CAG Val Lys Gln Ala Ile Ser Gly Thr Ser Pro Val Ala Ser Ser Phe Gln 485 490 495 | 1725 |
| GGT ATC AAA AAG TCT ATA TTC AGA ATA AGA GAA TCA CCT AAA GAG ACT Gly Ile Lys Lys Ser Ile Phe Arg Ile Arg Glu Ser Pro Lys Glu Thr 500 505 510 515 | 1773 |
| TTC AAT GCA AGT TTT TCA GGT CAT ATG ACT GAT CCA AAC TTT AAA AAA Phe Asn Ala Ser Phe Ser Gly His Met Thr Asp Pro Asn Phe Lys Lys 520 525 530 | 1821 |
| GAA ACT GAA GCC TCT GAA AGT GGA CTG GAA ATA CAT ACT GTT TGC TCA Glu Thr Glu Ala Ser Glu Ser Gly Leu Glu Ile His Thr Val Cys Ser 535 540 545 | 1869 |
| CAG AAG GAG GAC TCC TTA TGT CCA AAT TTA ATT GAT AAT GGA AGC TGG Gln Lys Glu Asp Ser Leu Cys Pro Asn Leu Ile Asp Asn Gly Ser Trp 550 555 560 | 1917 |
| CCA GCC ACC ACC ACA CAG AAT TCT GTA GCT TTG AAG AAT GCA GGT TTA Pro Ala Thr Thr Gln Asn Ser Val Ala Leu Lys Asn Ala Gly Leu 565 570 575 | 1965 |
| ATA TCC ACT TTG AAA AAG AAA ACA AAT AAG TTT ATT TAT GCT ATA CAT Ile Ser Thr Leu Lys Lys Thr Asn Lys Phe Ile Tyr Ala Ile His 580 585 590 595 | 2013 |
| GAT GAA ACA TCT TAT AAA GGA AAA AAA ATA CCG AAA GAC CAA AAA TCA Asp Glu Thr Ser Tyr Lys Gly Lys Lys Ile Pro Lys Asp Gln Lys Ser 600 605 610 | 2061 |
| GAA CTA ATT AAC TGT TCA GCC CAG TTT GAA GCA AAT GCT TTT GAA GCA Glu Leu Ile Asn Cys Ser Ala Gln Phe Glu Ala Asn Ala Phe Glu Ala 615 620 625 | 2109 |

D
D
CONK

| | |
|---|------|
| CCA CTT ACA TTT GCA AAT GCT GAT TCA GGT TTA TTG CAT TCT TCT GTG Pro Leu Thr Phe Ala Asn Ala Asp Ser Gly Leu Leu His Ser Ser Val 630 635 640 | 2157 |
| AAA AGA AGC TGT TCA CAG AAT GAT TCT GAA GAA CCA ACT TTG TCC TTA Lys Arg Ser Cys Ser Gln Asn Asp Ser Glu Glu Pro Thr Leu Ser Leu 645 650 655 | 2205 |
| ACT AGC TCT TTT GGG ACA ATT CTG AGG AAA TGT TCT AGA AAT GAA ACA Thr Ser Phe Gly Thr Ile Leu Arg Lys Cys Ser Arg Asn Glu Thr 660 665 670 675 | 2253 |
| TGT TCT AAT AAT ACA GTA ATC TCT CAG GAT CTT GAT TAT AAA GAA GCA Cys Ser Asn Asn Thr Val Ile Ser Gln Asp Leu Asp Tyr Lys Glu Ala 680 685 690 | 2301 |
| AAA TGT AAT AAG GAA AAA CTA CAG TTA TTT ATT ACC CCA GAA GCT GAT Lys Cys Asn Lys Glu Lys Leu Gln Leu Phe Ile Thr Pro Glu Ala Asp 695 700 705 | 2349 |
| TCT CTG TCA TGC CTG CAG GAA GGA CAG TGT GAA AAT GAT CCA AAA AGC Ser Leu Ser Cys Leu Gln Glu Gly Gln Cys Glu Asn Asp Pro Lys Ser 710 715 720 | 2397 |
| AAA AAA GTT TCA GAT ATA AAA GAA GAG GTC TTG GCT GCA GCA TGT CAC Lys Lys Val Ser Asp Ile Lys Glu Glu Val Leu Ala Ala Ala Cys His 725 730 735 | 2445 |
| CCA GTA CAA CAC TCA AAA GTG GAA TAC AGT GAT ACT GAC TTT CAA TCC Pro Val Gln His Ser Lys Val Glu Tyr Ser Asp Thr Asp Phe Gln Ser 740 745 750 755 | 2493 |
| CAG AAA AGT CTT TTA TAT GAT CAT GAA AAT GCC AGC ACT CTT ATT TTA Gln Lys Ser Leu Leu Tyr Asp His Glu Asn Ala Ser Thr Leu Ile Leu 760 765 770 | 2541 |
| ACT CCT ACT TCC AAG GAT GTT CTG TCA AAC CTA GTC ATG ATT TCT AGA Thr Pro Thr Ser Lys Asp Val Leu Ser Asn Leu Val Met Ile Ser Arg 775 780 785 | 2589 |
| GGC AAA GAA TCA TAC AAA ATG TCA GAC AAG CTC AAA GGT AAC AAT TAT Gly Lys Glu Ser Tyr Lys Met Ser Asp Lys Leu Lys Gly Asn Asn Tyr 790 795 800 | 2637 |
| GAA TCT GAT GTT GAA TTA ACC AAA AAT ATT CCC ATG GAA AAG AAT CAA Glu Ser Asp Val Glu Leu Thr Lys Asn Ile Pro Met Glu Lys Asn Gln 805 810 815 | 2685 |
| GAT GTA TGT GCT TTA AAT GAA AAT TAT AAA AAC GTT GAG CTG TTG CCA Asp Val Cys Ala Leu Asn Glu Asn Tyr Lys Asn Val Glu Leu Leu Pro 820 825 830 835 | 2733 |
| CCT GAA AAA TAC ATG AGA GTA GCA TCA CCT TCA AGA AAG GTA CAA TTC Pro Glu Lys Tyr Met Arg Val Ala Ser Pro Ser Arg Lys Val Gln Phe 840 845 850 | 2781 |
| AAC CAA AAC ACA AAT CTA AGA GTA ATC CAA AAA AAT CAA GAA GAA ACT Asn Gln Asn Thr Asn Leu Arg Val Ile Gln Lys Asn Gln Glu Glu Thr 855 860 865 | 2829 |
| ACT TCA ATT TCA AAA ATA ACT GTC AAT CCA GAC TCT GAA GAA CTT TTC | 2877 |

| | | | |
|---|------|------|------|
| Thr Ser Ile Ser Lys Ile Thr Val Asn Pro Asp Ser Glu Glu Leu Phe | | | |
| 870 | 875 | 880 | |
| TCA GAC AAT GAG AAT AAT TTT GTC TTC CAA ATA GCT AAT GAA AGG AAT | | | 2925 |
| Ser Asp Asn Glu Asn Asn Phe Val Phe Gln Ile Ala Asn Glu Arg Asn | | | |
| 885 | 890 | 895 | |
| AAT CTT GCT TTA GGA AAT ACT AAG GAA CTT CAT GAA ACA GAC TTG ACT | | | 2973 |
| Asn Leu Ala Leu Gly Asn Thr Lys Glu Leu His Glu Thr Asp Leu Thr | | | |
| 900 | 905 | 910 | 915 |
| TGT GTA AAC GAA CCC ATT TTC AAG AAC TCT ACC ATG GTT TTA TAT GGA | | | 3021 |
| Cys Val Asn Glu Pro Ile Phe Lys Asn Ser Thr Met Val Leu Tyr Gly | | | |
| 920 | 925 | 930 | |
| GAC ACA GGT GAT AAA CAA GCA ACC CAA GTG TCA ATT AAA AAA GAT TTG | | | 3069 |
| Asp Thr Gly Asp Lys Gln Ala Thr Gln Val Ser Ile Lys Lys Asp Leu | | | |
| 935 | 940 | 945 | |
| GTT TAT GTT CTT GCA GAG GAG AAC AAA AAT AGT GTA AAG CAG CAT ATA | | | 3117 |
| Val Tyr Val Leu Ala Glu Glu Asn Lys Asn Ser Val Lys Gln His Ile | | | |
| 950 | 955 | 960 | |
| AAA ATG ACT CTA GGT CAA GAT TTA AAA TCG GAC ATC TCC TTG AAT ATA | | | 3165 |
| Lys Met Thr Leu Gly Gln Asp Leu Lys Ser Asp Ile Ser Leu Asn Ile | | | |
| 965 | 970 | 975 | |
| GAT AAA ATA CCA GAA AAA AAT AAT GAT TAC ATG GAC AAA TGG GCA GGA | | | 3213 |
| Asp Lys Ile Pro Glu Lys Asn Asn Asp Tyr Met Asp Lys Trp Ala Gly | | | |
| 980 | 985 | 990 | 995 |
| CTC TTA GGT CCA ATT TCA AAT CAC AGT TTT GGA GGT AGC TTC AGA ACA | | | 3261 |
| Leu Leu Gly Pro Ile Ser Asn His Ser Phe Gly Gly Ser Phe Arg Thr | | | |
| 1000 | 1005 | 1010 | |
| GCT TCA AAT AAG GAA ATC AAG CTC TCT GAA CAT AAC ATT AAG AAG AGC | | | 3309 |
| Ala Ser Asn Lys Glu Ile Lys Leu Ser Glu His Asn Ile Lys Lys Ser | | | |
| 1015 | 1020 | 1025 | |
| AAA ATG TTC TTC AAA GAT ATT GAA GAA CAA TAT CCT ACT AGT TTA GCT | | | 3357 |
| Lys Met Phe Phe Lys Asp Ile Glu Glu Gln Tyr Pro Thr Ser Leu Ala | | | |
| 1030 | 1035 | 1040 | |
| TGT GTT GAA ATT GTA AAT ACC TTG GCA TTA GAT AAT CAA AAG AAA CTG | | | 3405 |
| Cys Val Glu Ile Val Asn Thr Leu Ala Leu Asp Asn Gln Lys Lys Leu | | | |
| 1045 | 1050 | 1055 | |
| AGC AAG CCT CAG TCA ATT AAT ACT GTA TCT GCA CAT TTA CAG AGT AGT | | | 3453 |
| Ser Lys Pro Gln Ser Ile Asn Thr Val Ser Ala His Leu Gln Ser Ser | | | |
| 1060 | 1065 | 1070 | 1075 |
| GTA GTT GTT TCT GAT TGT AAA AAT AGT CAT ATA ACC CCT CAG ATG TTA | | | 3501 |
| Val Val Val Ser Asp Cys Lys Asn Ser His Ile Thr Pro Gln Met Leu | | | |
| 1080 | 1085 | 1090 | |
| TTT TCC AAG CAG GAT TTT AAT TCA AAC CAT AAT TTA ACA CCT AGC CAA | | | 3549 |
| Phe Ser Lys Gln Asp Phe Asn Ser Asn His Asn Leu Thr Pro Ser Gln | | | |
| 1095 | 1100 | 1105 | |
| AAG GCA GAA ATT ACA GAA CTT TCT ACT ATA TTA GAA GAA TCA GGA AGT | | | 3597 |
| Lys Ala Glu Ile Thr Glu Leu Ser Thr Ile Leu Glu Glu Ser Gly Ser | | | |

| | | | |
|--|------|------|------|
| 1110 | 1115 | 1120 | |
| CAG TTT GAA TTT ACT CAG TTT AGA AAA CCA AGC TAC ATA TTG CAG AAG Gln Phe Phe Thr Gln Phe Arg Lys Pro Ser Tyr Ile Leu Gln Lys 1125 | 1130 | 1135 | 3645 |
| AGT ACA TTT GAA GTG CCT GAA AAC CAG ATG ACT ATC TTA AAG ACC ACT Ser Thr Phe Glu Val Pro Glu Asn Gln Met Thr Ile Leu Lys Thr Thr 1140 | 1145 | 1150 | 3693 |
| TCT GAG GAA TGC AGA GAT GCT GTC CTT CAT GTC ATA ATG AAT GCC CCA Ser Glu Glu Cys Arg Asp Ala Asp Leu His Val Ile Met Asn Ala Pro 1160 | 1165 | 1170 | 3741 |
| TCG ATT GGT CAG GTA GAC AGC AGC AAG CAA TTT GAA GGT ACA GTT GAA Ser Ile Gly Gln Val Asp Ser Ser Lys Gln Phe Glu Gly Thr Val Glu 1175 | 1180 | 1185 | 3789 |
| ATT AAA CGG AAG TTT GCT GGC CTG TTG AAA AAT GAC TGT AAC AAA AGT Ile Lys Arg Lys Phe Ala Gly Leu Leu Lys Asn Asp Cys Asn Lys Ser 1190 | 1195 | 1200 | 3837 |
| GCT TCT GGT TAT TTA ACA GAT GAA AAT GAA GTG GGG TTT AGG GGC TTT Ala Ser Gly Tyr Leu Thr Asp Glu Asn Glu Val Gly Phe Arg Gly Phe 1205 | 1210 | 1215 | 3885 |
| TAT TCT GCT CAT GGC ACA AAA CTG AAT GTT TCT ACT GAA GCT CTG CAA Tyr Ser Ala His Gly Thr Lys Leu Asn Val Ser Thr Glu Ala Leu Gln 1220 | 1225 | 1230 | 3933 |
| AAA GCT GTG AAA CTG TTT AGT GAT ATT GAG AAT ATT AGT GAG GAA ACT Lys Ala Val Lys Leu Phe Ser Asp Ile Glu Asn Ile Ser Glu Glu Thr 1240 | 1245 | 1250 | 3981 |
| TCT GCA GAG GTA CAT CCA ATA AGT TTA TCT TCA AGT AAA TGT CAT GAT Ser Ala Glu Val His Pro Ile Ser Leu Ser Ser Lys Cys His Asp 1255 | 1260 | 1265 | 4029 |
| TCT GTT GTT TCA ATG TTT AAG ATA GAA AAT CAT AAT GAT AAA ACT GTA Ser Val Val Ser Met Phe Lys Ile Glu Asn His Asn Asp Lys Thr Val 1270 | 1275 | 1280 | 4077 |
| AGT GAA AAA AAT AAT AAA TGC CAA CTG ATA TTA CAA AAT AAT ATT GAA Ser Glu Lys Asn Asn Lys Cys Gln Leu Ile Leu Gln Asn Asn Ile Glu 1285 | 1290 | 1295 | 4125 |
| ATG ACT ACT GGC ACT TTT GTT GAA GAA ATT ACT GAA AAT TAC AAG AGA Met Thr Thr Gly Thr Phe Val Glu Glu Ile Thr Glu Asn Tyr Lys Arg 1300 | 1305 | 1310 | 4173 |
| AAT ACT GAA AAT GAA GAT AAC AAA TAT ACT GCT GCC AGT AGA AAT TCT Asn Thr Glu Asn Glu Asp Asn Lys Tyr Thr Ala Ala Ser Arg Asn Ser 1320 | 1325 | 1330 | 4221 |
| CAT AAC TTA GAA TTT GAT GGC AGT GAT TCA AGT AAA AAT GAT ACT GTT His Asn Leu Glu Phe Asp Gly Ser Asp Ser Ser Lys Asn Asp Thr Val 1335 | 1340 | 1345 | 4269 |
| TGT ATT CAT AAA GAT GAA ACG GAC TTG CTA TTT ACT GAT CAG CAC AAC Cys Ile His Lys Asp Glu Thr Asp Leu Leu Phe Thr Asp Gln His Asn 1350 | 1355 | 1360 | 4317 |

ATA TGT CTT AAA TTA TCT GGC CAG TTT ATG AAG GAG GGA AAC ACT CAG 4365
Ile Cys Leu Lys Leu Ser Gly Gln Phe Met Lys Glu Gly Asn Thr Gln
1365 1370 1375

ATT AAA GAA GAT TTG TCA GAT TTA ACT TTT TTG GAA GTT GCG AAA GCT 4413
Ile Lys Glu Asp Leu Ser Asp Leu Thr Phe Leu Glu Val Ala Lys Ala
1380 1385 1390 1395

CAA GAA GCA TGT CAT GGT AAT ACT TCA AAT AAA GAA CAG TTA ACT GCT 4461
Gln Glu Ala Cys His Gly Asn Thr Ser Asn Lys Glu Gln Leu Thr Ala
1400 1405 1410

ACT AAA ACG GAG CAA AAT ATA AAA GAT TTT GAG ACT TCT GAT ACA TTT 4509
Thr Lys Thr Glu Gln Asn Ile Lys Asp Phe Glu Thr Ser Asp Thr Phe
1415 1420 1425

TTT CAG ACT GCA AGT GGG AAA AAT ATT AGT GTC GCC AAA GAG TCA TTT 4557
Phe Gln Thr Ala Ser Gly Lys Asn Ile Ser Val Ala Lys Glu Ser Phe
1430 1435 1440

AAT AAA ATT GTA AAT TTC TTT GAT CAG AAA CCA GAA GAA TTG CAT AAC 4605
Asn Lys Ile Val Asn Phe Phe Asp Gln Lys Pro Glu Glu Leu His Asn
1445 1450 1455

TTT TCC TTA AAT TCT GAA TTA CAT TCT GAC ATA AGA AAG AAC AAA ATG 4653
Phe Ser Leu Asn Ser Glu Leu His Ser Asp Ile Arg Lys Asn Lys Met
1460 1465 1470 1475

GAC ATT CTA AGT TAT GAG GAA ACA GAC ATA GTT AAA CAC AAA ATA CTG 4701
Asp Ile Leu Ser Tyr Glu Glu Thr Asp Ile Val Lys His Lys Ile Leu
1480 1485 1490

AAA GAA AGT GTC CCA GTT GGT ACT GGA AAT CAA CTA GTG ACC TTC CAG 4749
Lys Glu Ser Val Pro Val Gly Thr Gly Asn Gln Leu Val Thr Phe Gln
1495 1500 1505

GGA CAA CCC GAA CGT GAT GAA AAG ATC AAA GAA CCT ACT CTG TTG GGT 4797
Gly Gln Pro Glu Arg Asp Glu Lys Ile Lys Glu Pro Thr Leu Leu Gly
1510 1515 1520

TTT CAT ACA GCT AGC GGG AAA AAA GTT AAA ATT GCA AAG GAA TCT TTG 4845
Phe His Thr Ala Ser Gly Lys Val Lys Ile Ala Lys Glu Ser Leu
1525 1530 1535

GAC AAA GTG AAA AAC CTT TTT GAT GAA AAA GAG CAA GGT ACT AGT GAA 4893
Asp Lys Val Lys Asn Leu Phe Asp Glu Lys Glu Gln Gly Thr Ser Glu
1540 1545 1550 1555

ATC ACC AGT TTT AGC CAT CAA TGG GCA AAG ACC CTA AAG TAC AGA GAG 4941
Ile Thr Ser Phe Ser His Gln Trp Ala Lys Thr Leu Lys Tyr Arg Glu
1560 1565 1570

GCC TGT AAA GAC CTT GAA TTA GCA TGT GAG ACC ATT GAG ATC ACA GCT 4989
Ala Cys Lys Asp Leu Glu Leu Ala Cys Glu Thr Ile Glu Ile Thr Ala
1575 1580 1585

GCC CCA AAG TGT AAA GAA ATG CAG AAT TCT CTC AAT AAT GAT AAA AAC 5037
Ala Pro Lys Cys Lys Glu Met Gln Asn Ser Leu Asn Asn Asp Lys Asn
1590 1595 1600

D
W

| | |
|---|------|
| CTT GTT TCT ATT GAG ACT GTG GTG CCA CCT AAG CTC TTA AGT GAT AAT Leu Val Ser Ile Glu Thr Val Val Pro Pro Lys Leu Leu Ser Asp Asn 1605 1610 1615 | 5085 |
| TTA TGT AGA CAA ACT GAA AAT CTC AAA ACA TCA AAA AGT ATC TTT TTG Leu Cys Arg Gln Thr Glu Asn Leu Lys Thr Ser Lys Ser Ile Phe Leu 1620 1625 1630 1635 | 5133 |
| AAA GTT AAA GTA CAT GAA AAT GTA GAA AAA GAA ACA GCA AAA AGT CCT Lys Val Lys Val His Glu Asn Val Glu Lys Glu Thr Ala Lys Ser Pro 1640 1645 1650 | 5181 |
| GCA ACT TGT TAC ACA AAT CAG TCC CCT TAT TCA GTC ATT GAA AAT TCA Ala Thr Cys Tyr Thr Asn Gln Ser Pro Tyr Ser Val Ile Glu Asn Ser 1655 1660 1665 | 5229 |
| GCC TTA GCT TTT TAC ACA AGT TGT AGT AGA AAA ACT TCT GTG AGT CAG Ala Leu Ala Phe Tyr Thr Ser Cys Ser Arg Lys Thr Ser Val Ser Gln 1670 1675 1680 | 5277 |
| ACT TCA TTA CTT GAA GCA AAA AAA TGG CTT AGA GAA GGA ATA TTT GAT Thr Ser Leu Leu Glu Ala Lys Lys Trp Leu Arg Glu Gly Ile Phe Asp 1685 1690 1695 | 5325 |
| GGT CAA CCA GAA AGA ATA AAT ACT GCA GAT TAT GTA GGA AAT TAT TTG Gly Gln Pro Glu Arg Ile Asn Thr Ala Asp Tyr Val Gly Asn Tyr Leu 1700 1705 1710 1715 | 5373 |
| TAT GAA AAT AAT TCA AAC AGT ACT ATA GCT GAA AAT GAC AAA AAT CAT Tyr Glu Asn Asn Ser Asn Ser Thr Ile Ala Glu Asn Asp Lys Asn His 1720 1725 1730 | 5421 |
| CTC TCC GAA AAA CAA GAT ACT TAT TTA AGT AAC AGT AGC ATG TCT AAC Leu Ser Glu Lys Gln Asp Thr Tyr Leu Ser Asn Ser Met Ser Asn 1735 1740 1745 | 5469 |
| AGC TAT TCC TAC CAT TCT GAT GAG GTA TAT AAT GAT TCA GGA TAT CTC Ser Tyr Ser Tyr His Ser Asp Glu Val Tyr Asn Asp Ser Gly Tyr Leu 1750 1755 1760 | 5517 |
| TCA AAA AAT AAA CTT GAT TCT GGT ATT GAG CCA GTA TTG AAG AAT GTT Ser Lys Asn Lys Leu Asp Ser Gly Ile Glu Pro Val Leu Lys Asn Val 1765 1770 1775 | 5565 |
| GAA GAT CAA AAA AAC ACT AGT TTT TCC AAA GTA ATA TCC AAT GTA AAA Glu Asp Gln Lys Asn Thr Ser Phe Ser Lys Val Ile Ser Asn Val Lys 1780 1785 1790 1795 | 5613 |
| GAT GCA AAT GCA TAC CCA CAA ACT GTA AAT GAA GAT ATT TGC GTT GAG Asp Ala Asn Ala Tyr Pro Gln Thr Val Asn Glu Asp Ile Cys Val Glu 1800 1805 1810 | 5661 |
| GAA CTT GTG ACT AGC TCT TCA CCC TGC AAA AAT AAA AAT GCA GCC ATT Glu Leu Val Thr Ser Ser Pro Cys Lys Asn Lys Asn Ala Ala Ile 1815 1820 1825 | 5709 |
| AAA TTG TCC ATA TCT AAT AGT AAT AAT TTT GAG GTA GGG CCA CCT GCA Lys Leu Ser Ile Ser Asn Ser Asn Asn Phe Glu Val Gly Pro Pro Ala 1830 1835 1840 | 5757 |
| TTT AGG ATA GCC AGT GGT AAA ATC GTT TGT GTT TCA CAT GAA ACA ATT | 5805 |

Phe Arg Ile Ala Ser Gly Lys Ile Val Cys Val Ser His Glu Thr Ile
 1845 1850 1855

| | |
|---|------|
| AAA AAA GTG AAA GAC ATA TTT ACA GAC AGT TTC AGT AAA GTA ATT AAG | 5853 |
| Lys Lys Val Lys Asp Ile Phe Thr Asp Ser Phe Ser Lys Val Ile Lys | |
| 1860 1865 1870 1875 | |
| GAA AAC AAC GAG AAT AAA TCA AAA ATT TGC CAA ACG AAA ATT ATG GCA | 5901 |
| Glu Asn Asn Glu Asn Lys Ser Lys Ile Cys Gln Thr Lys Ile Met Ala | |
| 1880 1885 1890 | |
| GGT TGT TAC GAG GCA TTG GAT GAT TCA GAG GAT ATT CTT CAT AAC TCT | 5949 |
| Gly Cys Tyr Glu Ala Leu Asp Asp Ser Glu Asp Ile Leu His Asn Ser | |
| 1895 1900 1905 | |
| CTA GAT AAT GAT GAA TGT AGC ACG CAT TCA CAT AAG GTT TTT GCT GAC | 5997 |
| Leu Asp Asn Asp Glu Cys Ser Thr His Ser His Lys Val Phe Ala Asp | |
| 1910 1915 1920 | |
| ATT CAG AGT GAA GAA ATT TTA CAA CAT AAC CAA AAT ATG TCT GGA TTG | 6045 |
| Ile Gln Ser Glu Glu Ile Leu Gln His Asn Gln Asn Met Ser Gly Leu | |
| 1925 1930 1935 | |
| GAG AAA GTT TCT AAA ATA TCA CCT TGT GAT GTT AGT TTG GAA ACT TCA | 6093 |
| Glu Lys Val Ser Lys Ile Ser Pro Cys Asp Val Ser Leu Glu Thr Ser | |
| 1940 1945 1950 1955 | |
| GAT ATA TGT AAA TGT AGT ATA GGG AAG CTT CAT AAG TCA GTC TCA TCT | 6141 |
| Asp Ile Cys Lys Cys Ser Ile Gly Lys Leu His Lys Ser Val Ser Ser | |
| 1960 1965 1970 | |
| GCA AAT ACT TGT GGG ATT TTT AGC ACA GCA AGT GGA AAA TCT GTC CAG | 6189 |
| Ala Asn Thr Cys Gly Ile Phe Ser Thr Ala Ser Gly Lys Ser Val Gln | |
| 1975 1980 1985 | |
| GTA TCA GAT GCT TCA TTA CAA AAC GCA AGA CAA GTG TTT TCT GAA ATA | 6237 |
| Val Ser Asp Ala Ser Leu Gln Asn Ala Arg Gln Val Phe Ser Glu Ile | |
| 1990 1995 2000 | |
| GAA GAT AGT ACC AAG CAA GTC TTT TCC AAA GTA TTG TTT AAA AGT AAC | 6285 |
| Glu Asp Ser Thr Lys Gln Val Phe Ser Lys Val Leu Phe Lys Ser Asn | |
| 2005 2010 2015 | |
| GAA CAT TCA GAC CAG CTC ACA AGA GAA AAT ACT GCT ATA CGT ACT | 6333 |
| Glu His Ser Asp Gln Leu Thr Arg Glu Glu Asn Thr Ala Ile Arg Thr | |
| 2020 2025 2030 2035 | |
| CCA GAA CAT TTA ATA TCC CAA AAA GGC TTT TCA TAT AAT GTG GTA AAT | 6381 |
| Pro Glu His Leu Ile Ser Gln Lys Gly Phe Ser Tyr Asn Val Val Asn | |
| 2040 2045 2050 | |
| TCA TCT GCT TTC TCT GGA TTT AGT ACA GCA AGT GGA AAG CAA GTT TCC | 6429 |
| Ser Ser Ala Phe Ser Gly Phe Ser Thr Ala Ser Gly Lys Gln Val Ser | |
| 2055 2060 2065 | |
| ATT TTA GAA AGT TCC TTA CAC AAA GTT AAG GGA GTG TTA GAG GAA TTT | 6477 |
| Ile Leu Glu Ser Ser Leu His Lys Val Lys Gly Val Leu Glu Glu Phe | |
| 2070 2075 2080 | |
| GAT TTA ATC AGA ACT GAG CAT AGT CTT CAC TAT TCA CCT ACG TCT AGA | 6525 |
| Asp Leu Ile Arg Thr Glu His Ser Leu His Tyr Ser Pro Thr Ser Arg | |

2085 2090 2095

| | | | | |
|--|--|--|--|------|
| CAA AAT GTA TCA AAA ATA CTT CCT CGT GTT GAT AAG AGA AAC CCA GAG | | | | 6573 |
| Gln Asn Val Ser Lys Ile Leu Pro Arg Val Asp Lys Arg Asn Pro Glu | | | | |
| 2100 2105 2110 2115 | | | | |
| CAC TGT GTA AAC TCA GAA ATG GAA AAA ACC TGC AGT AAA GAA TTT AAA | | | | 6621 |
| His Cys Val Asn Ser Glu Met Glu Lys Thr Cys Ser Lys Glu Phe Lys | | | | |
| 2120 2125 2130 | | | | |
| TTA TCA AAT AAC TTA AAT GTT GAA GGT GGT TCT TCA GAA AAT AAT CAC | | | | 6669 |
| Leu Ser Asn Asn Leu Asn Val Glu Gly Ser Ser Glu Asn Asn His | | | | |
| 2135 2140 2145 | | | | |
| TCT ATT AAA GTT TCT CCA TAT CTC TCT CAA TTT CAA CAA GAC AAA CAA | | | | 6717 |
| Ser Ile Lys Val Ser Pro Tyr Leu Ser Gln Phe Gln Gln Asp Lys Gln | | | | |
| 2150 2155 2160 | | | | |
| CAG TTG GTA TTA GGA ACC AAA GTC TCA CTT GTT GAG AAC ATT CAT GTT | | | | 6765 |
| Gln Leu Val Leu Gly Thr Lys Val Ser Leu Val Glu Asn Ile His Val | | | | |
| 2165 2170 2175 | | | | |
| TTG GGA AAA GAA CAG GCT TCA CCT AAA AAC GTA AAA ATG GAA ATT GGT | | | | 6813 |
| Leu Gly Lys Glu Gln Ala Ser Pro Lys Asn Val Lys Met Glu Ile Gly | | | | |
| 2180 2185 2190 2195 | | | | |
| AAA ACT GAA ACT TTT TCT GAT GTT CCT GTG AAA ACA AAT ATA GAA GTT | | | | 6861 |
| Lys Thr Glu Thr Phe Ser Asp Val Pro Val Lys Thr Asn Ile Glu Val | | | | |
| 2200 2205 2210 | | | | |
| TGT TCT ACT TAC TCC AAA GAT TCA GAA AAC TAC TTT GAA ACA GAA GCA | | | | 6909 |
| Cys Ser Thr Tyr Ser Lys Asp Ser Glu Asn Tyr Phe Glu Thr Glu Ala | | | | |
| 2215 2220 2225 | | | | |
| GTA GAA ATT GCT AAA GCT TTT ATG GAA GAT GAT GAA CTG ACA GAT TCT | | | | 6957 |
| Val Glu Ile Ala Lys Ala Phe Met Glu Asp Asp Glu Leu Thr Asp Ser | | | | |
| 2230 2235 2240 | | | | |
| AAA CTG CCA AGT CAT GCC ACA CAT TCT CTT TTT ACA TGT CCC GAA AAT | | | | 7005 |
| Lys Leu Pro Ser His Ala Thr His Ser Leu Phe Thr Cys Pro Glu Asn | | | | |
| 2245 2250 2255 | | | | |
| GAG GAA ATG GTT TTG TCA AAT TCA AGA ATT GGA AAA AGA AGA GGA GAG | | | | 7053 |
| Glu Glu Met Val Leu Ser Asn Ser Arg Ile Gly Lys Arg Arg Gly Glu | | | | |
| 2260 2265 2270 2275 | | | | |
| CCC CTT ATC TTA GTG GGA GAA CCC TCA ATC AAA AGA AAC TTA TTA AAT | | | | 7101 |
| Pro Leu Ile Leu Val Gly Glu Pro Ser Ile Lys Arg Asn Leu Leu Asn | | | | |
| 2280 2285 2290 | | | | |
| GAA TTT GAC AGG ATA ATA GAA AAT CAA GAA AAA TCC TTA AAG GCT TCA | | | | 7149 |
| Glu Phe Asp Arg Ile Ile Glu Asn Gln Glu Lys Ser Leu Lys Ala Ser | | | | |
| 2295 2300 2305 | | | | |
| AAA AGC ACT CCA GAT GGC ACA ATA AAA GAT CGA AGA TTG TTT ATG CAT | | | | 7197 |
| Lys Ser Thr Pro Asp Gly Thr Ile Lys Asp Arg Arg Leu Phe Met His | | | | |
| 2310 2315 2320 | | | | |
| CAT GTT TCT TTA GAG CCG ATT ACC TGT GTA CCC TTT CGC ACA ACT AAG | | | | 7245 |
| His Val Ser Leu Glu Pro Ile Thr Cys Val Pro Phe Arg Thr Thr Lys | | | | |
| 2325 2330 2335 | | | | |

GAA CGT CAA GAG ATA CAG AAT CCA AAT TTT ACC GCA CCT GGT CAA GAA 7293
 Glu Arg Gln Glu Ile Gln Asn Pro Asn Phe Thr Ala Pro Gly Gln Glu
 2340 2345 2350 2355

TTT CTG TCT AAA TCT CAT TTG TAT GAA CAT CTG ACT TTG GAA AAA TCT 7341
 Phe Leu Ser Lys Ser His Leu Tyr Glu His Leu Thr Leu Glu Lys Ser
 2360 2365 2370

TCA AGC AAT TTA GCA GTT TCA GGA CAT CCA TTT TAT CAA GTT TCT GCT 7389
 Ser Ser Asn Leu Ala Val Ser Gly His Pro Phe Tyr Gln Val Ser Ala
 2375 2380 2385

ACA AGA AAT GAA AAA ATG AGA CAC TTG ATT ACT ACA GGC AGA CCA ACC 7437
 Thr Arg Asn Glu Lys Met Arg His Leu Ile Thr Thr Gly Arg Pro Thr
 2390 2395 2400

AAA GTC TTT GTT CCA CCT TTT AAA ACT AAA TCA CAT TTT CAC AGA GTT 7485
 Lys Val Phe Val Pro Pro Phe Lys Thr Lys Ser His Phe His Arg Val
 2405 2410 2415

GAA CAG TGT GTT AGG AAT ATT AAC TTG GAG GAA AAC AGA CAA AAG CAA 7533
 Glu Gln Cys Val Arg Asn Ile Asn Leu Glu Glu Asn Arg Gln Lys Gln
 2420 2425 2430 2435

AAC ATT GAT GGA CAT GGC TCT GAT GAT AGT AAA AAT AAG ATT AAT GAC 7581
 Asn Ile Asp Gly His Gly Ser Asp Asp Ser Lys Asn Lys Ile Asn Asp
 2440 2445 2450

AAT GAG ATT CAT CAG TTT AAC AAA AAC AAC TCC AAT CAA GCA GCA GCT 7629
 Asn Glu Ile His Gln Phe Asn Lys Asn Asn Ser Asn Gln Ala Ala Ala
 2455 2460 2465

GTA ACT TTC ACA AAG TGT GAA GAA GAA CCT TTA GAT TTA ATT ACA AGT 7677
 Val Thr Phe Thr Lys Cys Glu Glu Pro Leu Asp Leu Ile Thr Ser
 2470 2475 2480

CTT CAG AAT GCC AGA GAT ATA CAG GAT ATG CGA ATT AAG AAG AAA CAA 7725
 Leu Gln Asn Ala Arg Asp Ile Gln Asp Met Arg Ile Lys Lys Lys Gln
 2485 2490 2495

AGG CAA CGC GTC TTT CCA CAG CCA GGC AGT CTG TAT CTT GCA AAA ACA 7773
 Arg Gln Arg Val Phe Pro Gln Pro Gly Ser Leu Tyr Leu Ala Lys Thr
 2500 2505 2510 2515

TCC ACT CTG CCT CGA ATC TCT CTG AAA GCA GCA GTA GGA GGC CAA GTT 7821
 Ser Thr Leu Pro Arg Ile Ser Leu Lys Ala Ala Val Gly Gly Gln Val
 2520 2525 2530

CCC TCT GCG TGT TCT CAT AAA CAG CTG TAT ACG TAT GGC GTT TCT AAA 7869
 Pro Ser Ala Cys Ser His Lys Gln Leu Tyr Thr Tyr Gly Val Ser Lys
 2535 2540 2545

CAT TGC ATA AAA ATT AAC AGC AAA AAT GCA GAG TCT TTT CAG TTT CAC 7917
 His Cys Ile Lys Ile Asn Ser Lys Asn Ala Glu Ser Phe Gln Phe His
 2550 2555 2560

ACT GAA GAT TAT TTT GGT AAG GAA AGT TTA TGG ACT GGA AAA GGA ATA 7965
 Thr Glu Asp Tyr Phe Gly Lys Glu Ser Leu Trp Thr Gly Lys Gly Ile
 2565 2570 2575

| | |
|---|------|
| CAG TTG GCT GAT GGT GGA TGG CTC ATA CCC TCC AAT GAT GGA AAG GCT Gln Leu Ala Asp Gly Gly Trp Leu Ile Pro Ser Asn Asp Gly Lys Ala 2580 2585 2590 2595 | 8013 |
| GGA AAA GAA GAA TTT TAT AGG GCT CTG TGT GAC ACT CCA GGT GTG GAT Gly Lys Glu Glu Phe Tyr Arg Ala Leu Cys Asp Thr Pro Gly Val Asp 2600 2605 2610 | 8061 |
| CCA AAG CTT ATT TCT AGA ATT TGG GTT TAT AAT CAC TAT AGA TGG ATC Pro Lys Leu Ile Ser Arg Ile Trp Val Tyr Asn His Tyr Arg Trp Ile 2615 2620 2625 | 8109 |
| ATA TGG AAA CTG GCA GCT ATG GAA TGT GCC TTT CCT AAG GAA TTT GCT Ile Trp Lys Leu Ala Ala Met Glu Cys Ala Phe Pro Lys Glu Phe Ala 2630 2635 2640 | 8157 |
| AAT AGA TGC CTA AGC CCA GAA AGG GTG CTT CTT CAA CTA AAA TAC AGA Asn Arg Cys Leu Ser Pro Glu Arg Val Leu Leu Gln Leu Lys Tyr Arg 2645 2650 2655 | 8205 |
| TAT GAT ACG GAA ATT GAT AGA AGC AGA AGA TCG GCT ATA AAA AAG ATA Tyr Asp Thr Glu Ile Asp Arg Ser Arg Arg Ser Ala Ile Lys Lys Ile 2660 2665 2670 2675 | 8253 |
| ATG GAA AGG GAT GAC ACA GCT GCA AAA ACA CTT GTT CTC TGT GTT TCT Met Glu Arg Asp Asp Thr Ala Ala Lys Thr Leu Val Leu Cys Val Ser 2680 2685 2690 | 8301 |
| GAC ATA ATT TCA TTG AGC GCA AAT ATA TCT GAA ACT TCT AGC AAT AAA Asp Ile Ile Ser Leu Ser Ala Asn Ile Ser Glu Thr Ser Ser Asn Lys 2695 2700 2705 | 8349 |
| ACT AGT AGT GCA GAT ACC CAA AAA GTG GCC ATT ATT GAA CTT ACA GAT Thr Ser Ser Ala Asp Thr Gln Lys Val Ala Ile Ile Glu Leu Thr Asp 2710 2715 2720 | 8397 |
| GGG TGG TAT GCT GTT AAG GCC CAG TTA GAT CCT CCC CTC TTA GCT GTC Gly Trp Tyr Ala Val Lys Ala Gln Leu Asp Pro Pro Leu Leu Ala Val 2725 2730 2735 | 8445 |
| TTA AAG AAT GGC AGA CTG ACA GTT GGT CAG AAG ATT ATT CTT CAT GGA Leu Lys Asn Gly Arg Leu Thr Val Gly Gln Lys Ile Ile Leu His Gly 2740 2745 2750 2755 | 8493 |
| GCA GAA CTG GTG GGC TCT CCT GAT GCC TGT ACA CCT CTT GAA GCC CCA Ala Glu Leu Val Gly Ser Pro Asp Ala Cys Thr Pro Leu Glu Ala Pro 2760 2765 2770 | 8541 |
| GAA TCT CTT ATG TTA AAG ATT TCT GCT AAC AGT ACT CGG CCT GCT CGC Glu Ser Leu Met Leu Lys Ile Ser Ala Asn Ser Thr Arg Pro Ala Arg 2775 2780 2785 | 8589 |
| TGG TAT ACC AAA CTT GGA TTC TTT CCT GAC CCT AGA CCT TTT CCT CTG Trp Tyr Thr Lys Leu Gly Phe Pro Asp Pro Arg Pro Phe Pro Leu 2790 2795 2800 | 8637 |
| CCC TTA TCA TCG CTT TTC AGT GAT GGA GGA AAT GTT GGT TGT GTT GAT Pro Leu Ser Ser Leu Phe Ser Asp Gly Gly Asn Val Gly Cys Val Asp 2805 2810 2815 | 8685 |
| GTA ATT ATT CAA AGA GCA TAC CCT ATA CAG TGG ATG GAG AAG ACA TCA | 8733 |

Val Ile Ile Gln Arg Ala Tyr Pro Ile Gln Trp Met Glu Lys Thr Ser
 2820 2825 2830 2835

TCT GGA TTA TAC ATA TTT CGC AAT GAA AGA GAG GAA GAA AAG GAA GCA
 Ser Gly Leu Tyr Ile Phe Arg Asn Glu Arg Glu Glu Lys Glu Ala
 2840 2845 2850 8781

GCA AAA TAT GTG GAG GCC CAA CAA AAG AGA CTA GAA GCC TTA TTC ACT
 Ala Lys Tyr Val Glu Ala Gln Gln Lys Arg Leu Glu Ala Leu Phe Thr
 2855 2860 2865 8829

AAA ATT CAG GAG GAA TTT GAA GAA CAT GAA GAA AAC ACA ACA AAA CCA
 Lys Ile Gln Glu Glu Phe Glu Glu His Glu Glu Asn Thr Thr Lys Pro
 2870 2875 2880 8877

TAT TTA CCA TCA CGT GCA CTA ACA AGA CAG CAA GTT CGT GCT TTG CAA
 Tyr Leu Pro Ser Arg Ala Leu Thr Arg Gln Gln Val Arg Ala Leu Gln
 2885 2890 2895 8925

GAT GGT GCA GAG CTT TAT GAA GCA GTG AAG AAT GCA GCA GAC CCA GCT
 Asp Gly Ala Glu Leu Tyr Glu Ala Val Lys Asn Ala Ala Asp Pro Ala
 2900 2905 2910 2915 8973


 TAC CTT GAG GGT TAT TTC AGT GAA GAG CAG TTA AGA GCC TTG AAT AAT
 Tyr Leu Glu Gly Tyr Phe Ser Glu Glu Gln Leu Arg Ala Leu Asn Asn
 2920 2925 2930 9021

CAC AGG CAA ATG TTG AAT GAT AAG AAA CAA GCT CAG ATC CAG TTG GAA
 His Arg Gln Met Leu Asn Asp Lys Lys Gln Ala Gln Ile Gln Leu Glu
 2935 2940 2945 9069

ATT AGG AAG ACC ATG GAA TCT GCT GAA CAA AAG GAA CAA GGT TTA TCA
 Ile Arg Lys Thr Met Glu Ser Ala Glu Gln Lys Glu Gln Gly Leu Ser
 2950 2955 2960 9117

AGG GAT GTC ACA ACC GTG TGG AAG TTG CGT ATT GTA AGC TAT TCA AAA
 Arg Asp Val Thr Thr Val Trp Lys Leu Arg Ile Val Ser Tyr Ser Lys
 2965 2970 2975 9165

AAA GAA AAA GAT TCA GTT ATA CTG AGT ATT TGG CGT CCA TCA TCA GAT
 Lys Glu Lys Asp Ser Val Ile Leu Ser Ile Trp Arg Pro Ser Ser Asp
 2980 2985 2990 2995 9213

TTA TAT TCT CTG TTA ACA GAA GGA AAG AGA TAC AGA ATT TAT CAT CTT
 Leu Tyr Ser Leu Leu Thr Glu Gly Lys Arg Tyr Arg Ile Tyr His Leu
 3000 3005 3010 9261

GCA ACT TCA AAA TCT AAA AGT AAA TCT GAA AGA GCT AAC ATA CAG TTA
 Ala Thr Ser Lys Ser Lys Ser Glu Arg Ala Asn Ile Gln Leu
 3015 3020 3025 9309

GCA GCG ACA AAA AAA ACT CAG TAT CAA CAA CTA CCG GTT TCA GAT GAA
 Ala Ala Thr Lys Lys Thr Gln Tyr Gln Gln Leu Pro Val Ser Asp Glu
 3030 3035 3040 9357

ATT TTA TTT CAG ATT TAC CAG CCA CGG GAG CCC CTT CAC TTC AGC AAA
 Ile Leu Phe Gln Ile Tyr Gln Pro Arg Glu Pro Leu His Phe Ser Lys
 3045 3050 3055 9405

TTT TTA GAT CCA GAC TTT CAG CCA TCT TGT TCT GAG GTG GAC CTA ATA
 Phe Leu Asp Pro Asp Phe Gln Pro Ser Cys Ser Glu Val Asp Leu Ile
 9453

| 3060 | 3065 | 3070 | 3075 | |
|--|------|------|------|-------|
| GGA TTT GTC GTT TCT GTG AAA AAA ACA GGA CTT GCC CCT TTC GTC Gly Phe Val Val Ser Val Val Lys Lys Thr Gly Leu Ala Pro Phe Val | | | | 9501 |
| 3080 | 3085 | | 3090 | |
| TAT TTG TCA GAC GAA TGT TAC AAT TTA CTG GCA ATA AAG TTT TGG ATA Tyr Leu Ser Asp Glu Cys Tyr Asn Leu Leu Ala Ile Lys Phe Trp Ile | | | | 9549 |
| 3095 | 3100 | | 3105 | |
| GAC CTT AAT GAG GAC ATT ATT AAG CCT CAT ATG TTA ATT GCT GCA AGC Asp Leu Asn Glu Asp Ile Ile Lys Pro His Met Leu Ile Ala Ala Ser | | | | 9597 |
| 3110 | 3115 | | 3120 | |
| AAC CTC CAG TGG CGA CCA GAA TCC AAA TCA GGC CTT CTT ACT TTA TTT Asn Leu Gln Trp Arg Pro Glu Ser Lys Ser Gly Leu Leu Thr Leu Phe | | | | 9645 |
| 3125 | 3130 | | 3135 | |
| GCT GGA GAT TTT TCT GTG TTT TCT GCT AGT CCA AAA GAG GGC CAC TTT Ala Gly Asp Phe Ser Val Phe Ser Ala Ser Pro Lys Glu Gly His Phe | | | | 9693 |
| 3140 | 3145 | | 3150 | |
| CAA GAG ACA TTC AAC AAA ATG AAA AAT ACT GTT GAG AAT ATT GAC ATA Gln Glu Thr Phe Asn Lys Met Lys Asn Thr Val Glu Asn Ile Asp Ile | | | | 9741 |
| 3160 | 3165 | | 3170 | |
| CTT TGC AAT GAA GCA GAA AAC AAG CTT ATG CAT ATA CTG CAT GCA AAT Leu Cys Asn Glu Ala Glu Asn Lys Leu Met His Ile Leu His Ala Asn | | | | 9789 |
| 3175 | 3180 | | 3185 | |
| GAT CCC AAG TGG TCC ACC CCA ACT AAA GAC TGT ACT TCA GGG CCG TAC Asp Pro Lys Trp Ser Thr Pro Thr Lys Asp Cys Thr Ser Gly Pro Tyr | | | | 9837 |
| 3190 | 3195 | | 3200 | |
| ACT GCT CAA ATC ATT CCT GGT ACA GGA AAC AAG CTT CTG ATG TCT TCT Thr Ala Gln Ile Ile Pro Gly Thr Gly Asn Lys Leu Leu Met Ser Ser | | | | 9885 |
| 3205 | 3210 | | 3215 | |
| CCT AAT TGT GAG ATA TAT TAT CAA AGT CCT TTA TCA CTT TGT ATG GCC Pro Asn Cys Glu Ile Tyr Tyr Gln Ser Pro Leu Ser Leu Cys Met Ala | | | | 9933 |
| 3220 | 3225 | | 3230 | |
| AAA AGG AAG TCT GTT TCC ACA CCT GTC TCA GCC CAG ATG ACT TCA AAG Lys Arg Lys Ser Val Ser Thr Pro Val Ser Ala Gln Met Thr Ser Lys | | | | 9981 |
| 3240 | 3245 | | 3250 | |
| TCT TGT AAA GGG GAG AAA GAG ATT GAT GAC CAA AAG AAC TGC AAA AAG Ser Cys Lys Gly Glu Lys Glu Ile Asp Asp Gln Lys Asn Cys Lys Lys | | | | 10029 |
| 3255 | 3260 | | 3265 | |
| AGA AGA GCC TTG GAT TTC TTG AGT AGA CTG CCT TTA CCT CCA CCT GTT Arg Arg Ala Leu Asp Phe Leu Ser Arg Leu Pro Leu Pro Pro Val | | | | 10077 |
| 3270 | 3275 | | 3280 | |
| AGT CCC ATT TGT ACA TTT GTT TCT CCG GCT GCA CAG AAG GCA TTT CAG Ser Pro Ile Cys Thr Phe Val Ser Pro Ala Ala Gln Lys Ala Phe Gln | | | | 10125 |
| 3285 | 3290 | | 3295 | |
| CCA CCA AGG AGT TGT GGC ACC AAA TAC GAA ACA CCC ATA AAG AAA AAA Pro Pro Arg Ser Cys Gly Thr Lys Tyr Glu Thr Pro Ile Lys Lys Lys | | | | 10173 |
| 3300 | 3305 | | 3310 | |
| | | | 3315 | |

GAA CTG AAT TCT CCT CAG ATG ACT CCA TTT AAA AAA TTC AAT GAA ATT 10221
 Glu Leu Asn Ser Pro Gln Met Thr Pro Phe Lys Lys Phe Asn Glu Ile
 3320 3325 3330

 TCT CTT TTG GAA AGT AAT TCA ATA GCT GAC GAA GAA CTT GCA TTG ATA 10269
 Ser Leu Leu Glu Ser Asn Ser Ile Ala Asp Glu Glu Leu Ala Leu Ile
 3335 3340 3345

 AAT ACC CAA GCT CTT TTG TCT GGT TCA ACA GGA GAA AAA CAA TTT ATA 10317
 Asn Thr Gln Ala Leu Leu Ser Gly Ser Thr Gly Glu Lys Gln Phe Ile
 3350 3355 3360

 TCT GTC AGT GAA TCC ACT AGG ACT GCT CCC ACC AGT TCA GAA GAT TAT 10365
 Ser Val Ser Glu Ser Thr Arg Thr Ala Pro Thr Ser Ser Glu Asp Tyr
 3365 3370 3375

 CTC AGA CTG AAA CGA CGT TGT ACT ACA TCT CTG ATC AAA GAA CAG GAG 10413
 Leu Arg Leu Lys Arg Arg Cys Thr Thr Ser Leu Ile Lys Glu Gln Glu
 3380 3385 3390 3395

 AGT TCC CAG GCC AGT ACG GAA GAA TGT GAG AAA AAT AAG CAG GAC ACA 10461
 Ser Ser Gln Ala Ser Thr Glu Glu Cys Glu Lys Asn Lys Gln Asp Thr
 3400 3405 3410

 ATT ACA ACT AAA AAA TAT ATC TAA 10485
 Ile Thr Thr Lys Lys Tyr Ile
 3415

D X CMK

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3418 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Met Pro Ile Gly Ser Lys Glu Arg Pro Thr Phe Phe Glu Ile Phe Lys
 1 5 10 15
 Thr Arg Cys Asn Lys Ala Asp Leu Gly Pro Ile Ser Leu Asn Trp Phe
 20 25 30
 Glu Glu Leu Ser Ser Glu Ala Pro Pro Tyr Asn Ser Glu Pro Ala Glu
 35 40 45
 Glu Ser Glu His Lys Asn Asn Tyr Glu Pro Asn Leu Phe Lys Thr
 50 55 60
 Pro Gln Arg Lys Pro Ser Tyr Asn Gln Leu Ala Ser Thr Pro Ile Ile
 65 70 75 80
 Phe Lys Glu Gln Gly Leu Thr Leu Pro Leu Tyr Gln Ser Pro Val Lys
 85 90 95
 Glu Leu Asp Lys Phe Lys Leu Asp Leu Gly Arg Asn Val Pro Asn Ser
 100 105 110
 Arg His Lys Ser Leu Arg Thr Val Lys Thr Lys Met Asp Gln Ala Asp
 115 120 125
 Asp Val Ser Cys Pro Leu Leu Asn Ser Cys Leu Ser Glu Ser Pro Val
 130 135 140

Val Leu Gln Cys Thr His Val Thr Pro Gln Arg Asp Lys Ser Val Val
145 150 155 160
Cys Gly Ser Leu Phe His Thr Pro Lys Phe Val Lys Gly Arg Gln Thr
165 170 175
Pro Lys His Ile Ser Glu Ser Leu Gly Ala Glu Val Asp Pro Asp Met
180 185 190
Ser Trp Ser Ser Ser Leu Ala Thr Pro Pro Thr Leu Ser Ser Thr Val
195 200 205
Leu Ile Val Arg Asn Glu Glu Ala Ser Glu Thr Val Phe Pro His Asp
210 215 220
Thr Thr Ala Asn Val Lys Ser Tyr Phe Ser Asn His Asp Glu Ser Leu
225 230 235 240
Lys Lys Asn Asp Arg Phe Ile Ala Ser Val Thr Asp Ser Glu Asn Thr
245 250 255
Asn Gln Arg Glu Ala Ala Ser His Gly Phe Gly Lys Thr Ser Gly Asn
260 265 270
Ser Phe Lys Val Asn Ser Cys Lys Asp His Ile Gly Lys Ser Met Pro
275 280 285
His Val Leu Glu Asp Glu Val Tyr Glu Thr Val Val Asp Thr Ser Glu
290 295 300
Glu Asp Ser Phe Ser Leu Cys Phe Ser Lys Cys Arg Thr Lys Asn Leu
305 310 315 320
Gln Lys Val Arg Thr Ser Lys Thr Arg Lys Lys Ile Phe His Glu Ala
325 330 335
Asn Ala Asp Glu Cys Glu Lys Ser Lys Asn Gln Val Lys Glu Lys Tyr
340 345 350
Ser Phe Val Ser Glu Val Glu Pro Asn Asp Thr Asp Pro Leu Asp Ser
355 360 365
Asn Val Ala His Gln Lys Pro Phe Glu Ser Gly Ser Asp Lys Ile Ser
370 375 380
Lys Glu Val Val Pro Ser Leu Ala Cys Glu Trp Ser Gln Leu Thr Leu
385 390 395 400
Ser Gly Leu Asn Gly Ala Gln Met Glu Lys Ile Pro Leu Leu His Ile
405 410 415
Ser Ser Cys Asp Gln Asn Ile Ser Glu Lys Asp Leu Leu Asp Thr Glu
420 425 430
Asn Lys Arg Lys Lys Asp Phe Leu Thr Ser Glu Asn Ser Leu Pro Arg
435 440 445
Ile Ser Ser Leu Pro Lys Ser Glu Lys Pro Leu Asn Glu Glu Thr Val
450 455 460
Val Asn Lys Arg Asp Glu Glu Gln His Leu Glu Ser His Thr Asp Cys
465 470 475 480
Ile Leu Ala Val Lys Gln Ala Ile Ser Gly Thr Ser Pro Val Ala Ser
485 490 495
Ser Phe Gln Gly Ile Lys Lys Ser Ile Phe Arg Ile Arg Glu Ser Pro
500 505 510
Lys Glu Thr Phe Asn Ala Ser Phe Ser Gly His Met Thr Asp Pro Asn
515 520 525
Phe Lys Lys Glu Thr Glu Ala Ser Glu Ser Gly Leu Glu Ile His Thr
530 535 540
Val Cys Ser Gln Lys Glu Asp Ser Leu Cys Pro Asn Leu Ile Asp Asn
545 550 555 560
Gly Ser Trp Pro Ala Thr Thr Gln Asn Ser Val Ala Leu Lys Asn
565 570 575
Ala Gly Leu Ile Ser Thr Leu Lys Lys Thr Asn Lys Phe Ile Tyr
580 585 590
Ala Ile His Asp Glu Thr Ser Tyr Lys Gly Lys Lys Ile Pro Lys Asp
595 600 605
Gln Lys Ser Glu Leu Ile Asn Cys Ser Ala Gln Phe Glu Ala Asn Ala
610 615 620
Phe Glu Ala Pro Leu Thr Phe Ala Asn Ala Asp Ser Gly Leu Leu His

625 630 635 640
Ser Ser Val Lys Arg Ser Cys Ser Gln Asn Asp Ser Glu Glu Pro Thr
645 650 655
Leu Ser Leu Thr Ser Ser Phe Gly Thr Ile Leu Arg Lys Cys Ser Arg
660 665 670
Asn Glu Thr Cys Ser Asn Asn Thr Val Ile Ser Gln Asp Leu Asp Tyr
675 680 685
Lys Glu Ala Lys Cys Asn Lys Glu Lys Leu Gln Leu Phe Ile Thr Pro
690 695 700
Glu Ala Asp Ser Leu Ser Cys Leu Gln Glu Gly Gln Cys Glu Asn Asp
705 710 715 720
Pro Lys Ser Lys Lys Val Ser Asp Ile Lys Glu Glu Val Leu Ala Ala
725 730 735
Ala Cys His Pro Val Gln His Ser Lys Val Glu Tyr Ser Asp Thr Asp
740 745 750
Phe Gln Ser Gln Lys Ser Leu Leu Tyr Asp His Glu Asn Ala Ser Thr
755 760 765
Leu Ile Leu Thr Pro Thr Ser Lys Asp Val Leu Ser Asn Leu Val Met
770 775 780
Ile Ser Arg Gly Lys Glu Ser Tyr Lys Met Ser Asp Lys Leu Lys Gly
785 790 795 800
Asn Asn Tyr Glu Ser Asp Val Glu Leu Thr Lys Asn Ile Pro Met Glu
805 810 815
Lys Asn Gln Asp Val Cys Ala Leu Asn Glu Asn Tyr Lys Asn Val Glu
820 825 830
Leu Leu Pro Pro Glu Lys Tyr Met Arg Val Ala Ser Pro Ser Arg Lys
835 840 845
Val Gln Phe Asn Gln Asn Thr Asn Leu Arg Val Ile Gln Lys Asn Gln
850 855 860
Glu Glu Thr Thr Ser Ile Ser Lys Ile Thr Val Asn Pro Asp Ser Glu
865 870 875 880
Glu Leu Phe Ser Asp Asn Glu Asn Asn Phe Val Phe Gln Ile Ala Asn
885 890 895
Glu Arg Asn Asn Leu Ala Leu Gly Asn Thr Lys Glu Leu His Glu Thr
900 905 910
Asp Leu Thr Cys Val Asn Glu Pro Ile Phe Lys Asn Ser Thr Met Val
915 920 925
Leu Tyr Gly Asp Thr Gly Asp Lys Gln Ala Thr Gln Val Ser Ile Lys
930 935 940
Lys Asp Leu Val Tyr Val Leu Ala Glu Glu Asn Lys Asn Ser Val Lys
945 950 955 960
Gln His Ile Lys Met Thr Leu Gly Gln Asp Leu Lys Ser Asp Ile Ser
965 970 975
Leu Asn Ile Asp Lys Ile Pro Glu Lys Asn Asn Asp Tyr Met Asp Lys
980 985 990
Trp Ala Gly Leu Leu Gly Pro Ile Ser Asn His Ser Phe Gly Gly Ser
995 1000 1005
Phe Arg Thr Ala Ser Asn Lys Glu Ile Lys Leu Ser Glu His Asn Ile
1010 1015 1020
Lys Lys Ser Lys Met Phe Phe Lys Asp Ile Glu Glu Gln Tyr Pro Thr
1025 1030 1035 1040
Ser Leu Ala Cys Val Glu Ile Val Asn Thr Leu Ala Leu Asp Asn Gln
1045 1050 1055
Lys Lys Leu Ser Lys Pro Gln Ser Ile Asn Thr Val Ser Ala His Leu
1060 1065 1070
Gln Ser Ser Val Val Val Ser Asp Cys Lys Asn Ser His Ile Thr Pro
1075 1080 1085
Gln Met Leu Phe Ser Lys Gln Asp Phe Asn Ser Asn His Asn Leu Thr
1090 1095 1100
Pro Ser Gln Lys Ala Glu Ile Thr Glu Leu Ser Thr Ile Leu Glu Glu
1105 1110 1115 112

Ser Gly Ser Gln Phe Glu Phe Thr Gln Phe Arg Lys Pro Ser Tyr Ile
1125 1130 1135
Leu Gln Lys Ser Thr Phe Glu Val Pro Glu Asn Gln Met Thr Ile Leu
1140 1145 1150
Lys Thr Thr Ser Glu Glu Cys Arg Asp Ala Asp Leu His Val Ile Met
1155 1160 1165
Asn Ala Pro Ser Ile Gly Gln Val Asp Ser Ser Lys Gln Phe Glu Gly
1170 1175 1180
Thr Val Glu Ile Lys Arg Lys Phe Ala Gly Leu Leu Lys Asn Asp Cys
1185 1190 1195 1200
Asn Lys Ser Ala Ser Gly Tyr Leu Thr Asp Glu Asn Glu Val Gly Phe
1205 1210 1215
Arg Gly Phe Tyr Ser Ala His Gly Thr Lys Leu Asn Val Ser Thr Glu
1220 1225 1230
Ala Leu Gln Lys Ala Val Lys Leu Phe Ser Asp Ile Glu Asn Ile Ser
1235 1240 1245
Glu Glu Thr Ser Ala Glu Val His Pro Ile Ser Leu Ser Ser Ser Lys
1250 1255 1260
Cys His Asp Ser Val Val Ser Met Phe Lys Ile Glu Asn His Asn Asp
1265 1270 1275 1280
Lys Thr Val Ser Glu Lys Asn Asn Lys Cys Gln Leu Ile Leu Gln Asn
1285 1290 1295
Asn Ile Glu Met Thr Thr Gly Thr Phe Val Glu Glu Ile Thr Glu Asn
1300 1305 1310
Tyr Lys Arg Asn Thr Glu Asn Glu Asp Asn Lys Tyr Thr Ala Ala Ser
1315 1320 1325
Arg Asn Ser His Asn Leu Glu Phe Asp Gly Ser Asp Ser Ser Lys Asn
1330 1335 1340
Asp Thr Val Cys Ile His Lys Asp Glu Thr Asp Leu Leu Phe Thr Asp
1345 1350 1355 1360
Gln His Asn Ile Cys Leu Lys Leu Ser Gly Gln Phe Met Lys Glu Gly
1365 1370 1375
Asn Thr Gln Ile Lys Glu Asp Leu Ser Asp Leu Thr Phe Leu Glu Val
1380 1385 1390
Ala Lys Ala Gln Glu Ala Cys His Gly Asn Thr Ser Asn Lys Glu Gln
1395 1400 1405
Leu Thr Ala Thr Lys Thr Glu Gln Asn Ile Lys Asp Phe Glu Thr Ser
1410 1415 1420
Asp Thr Phe Phe Gln Thr Ala Ser Gly Lys Asn Ile Ser Val Ala Lys
1425 1430 1435 1440
Glu Ser Phe Asn Lys Ile Val Asn Phe Phe Asp Gln Lys Pro Glu Glu
1445 1450 1455
Leu His Asn Phe Ser Leu Asn Ser Glu Leu His Ser Asp Ile Arg Lys
1460 1465 1470
Asn Lys Met Asp Ile Leu Ser Tyr Glu Glu Thr Asp Ile Val Lys His
1475 1480 1485
Lys Ile Leu Lys Glu Ser Val Pro Val Gly Thr Gly Asn Gln Leu Val
1490 1495 1500
Thr Phe Gln Gly Gln Pro Glu Arg Asp Glu Lys Ile Lys Glu Pro Thr
1505 1510 1515 1520
Leu Leu Gly Phe His Thr Ala Ser Gly Lys Lys Val Lys Ile Ala Lys
1525 1530 1535
Glu Ser Leu Asp Lys Val Lys Asn Leu Phe Asp Glu Lys Glu Gln Gly
1540 1545 1550
Thr Ser Glu Ile Thr Ser Phe Ser His Gln Trp Ala Lys Thr Leu Lys
1555 1560 1565
Tyr Arg Glu Ala Cys Lys Asp Leu Glu Leu Ala Cys Glu Thr Ile Glu
1570 1575 1580
Ile Thr Ala Ala Pro Lys Cys Lys Glu Met Gln Asn Ser Leu Asn Asn
1585 1590 1595 1600
Asp Lys Asn Leu Val Ser Ile Glu Thr Val Val Pro Pro Lys Leu Leu

1605 1610 1615
 Ser Asp Asn Leu Cys Arg Gln Thr Glu Asn Leu Lys Thr Ser Lys Ser
 1620 1625 1630
 Ile Phe Leu Lys Val Lys Val His Glu Asn Val Glu Lys Glu Thr Ala
 1635 1640 1645
 Lys Ser Pro Ala Thr Cys Tyr Thr Asn Gln Ser Pro Tyr Ser Val Ile
 1650 1655 1660
 Glu Asn Ser Ala Leu Ala Phe Tyr Thr Ser Cys Ser Arg Lys Thr Ser
 1665 1670 1675 1680
 Val Ser Gln Thr Ser Leu Leu Glu Ala Lys Lys Trp Leu Arg Glu Gly
 1685 1690 1695
 Ile Phe Asp Gly Gln Pro Glu Arg Ile Asn Thr Ala Asp Tyr Val Gly
 1700 1705 1710
 Asn Tyr Leu Tyr Glu Asn Asn Ser Asn Ser Thr Ile Ala Glu Asn Asp
 1715 1720 1725
 Lys Asn His Leu Ser Glu Lys Gln Asp Thr Tyr Leu Ser Asn Ser Ser
 1730 1735 1740
 Met Ser Asn Ser Tyr Ser Tyr His Ser Asp Glu Val Tyr Asn Asp Ser
 1745 1750 1755 1760
 Gly Tyr Leu Ser Lys Asn Lys Leu Asp Ser Gly Ile Glu Pro Val Leu
 1765 1770 1775
 Lys Asn Val Glu Asp Gln Lys Asn Thr Ser Phe Ser Lys Val Ile Ser
 1780 1785 1790
 Asn Val Lys Asp Ala Asn Ala Tyr Pro Gln Thr Val Asn Glu Asp Ile
 1795 1800 1805
 Cys Val Glu Glu Leu Val Thr Ser Ser Ser Pro Cys Lys Asn Lys Asn
 1810 1815 1820
 Ala Ala Ile Lys Leu Ser Ile Ser Asn Ser Asn Asn Phe Glu Val Gly
 1825 1830 1835 1840
 Pro Pro Ala Phe Arg Ile Ala Ser Gly Lys Ile Val Cys Val Ser His
 1845 1850 1855
 Glu Thr Ile Lys Lys Val Lys Asp Ile Phe Thr Asp Ser Phe Ser Lys
 1860 1865 1870
 Val Ile Lys Glu Asn Asn Glu Asn Lys Ser Lys Ile Cys Gln Thr Lys
 1875 1880 1885
 Ile Met Ala Gly Cys Tyr Glu Ala Leu Asp Asp Ser Glu Asp Ile Leu
 1890 1895 1900
 His Asn Ser Leu Asp Asn Asp Glu Cys Ser Thr His Ser His Lys Val
 1905 1910 1915 1920
 Phe Ala Asp Ile Gln Ser Glu Glu Ile Leu Gln His Asn Gln Asn Met
 1925 1930 1935
 Ser Gly Leu Glu Lys Val Ser Lys Ile Ser Pro Cys Asp Val Ser Leu
 1940 1945 1950
 Glu Thr Ser Asp Ile Cys Lys Cys Ser Ile Gly Lys Leu His Lys Ser
 1955 1960 1965
 Val Ser Ser Ala Asn Thr Cys Gly Ile Phe Ser Thr Ala Ser Gly Lys
 1970 1975 1980
 Ser Val Gln Val Ser Asp Ala Ser Leu Gln Asn Ala Arg Gln Val Phe
 1985 1990 1995 2000
 Ser Glu Ile Glu Asp Ser Thr Lys Gln Val Phe Ser Lys Val Leu Phe
 2005 2010 2015
 Lys Ser Asn Glu His Ser Asp Gln Leu Thr Arg Glu Glu Asn Thr Ala
 2020 2025 2030
 Ile Arg Thr Pro Glu His Leu Ile Ser Gln Lys Gly Phe Ser Tyr Asn
 2035 2040 2045
 Val Val Asn Ser Ser Ala Phe Ser Gly Phe Ser Thr Ala Ser Gly Lys
 2050 2055 2060
 Gln Val Ser Ile Leu Glu Ser Ser Leu His Lys Val Lys Gly Val Leu
 2065 2070 2075 2080
 Glu Glu Phe Asp Leu Ile Arg Thr Glu His Ser Leu His Tyr Ser Pro
 2085 2090 2095

Thr Ser Arg Gln Asn Val Ser Lys Ile Leu Pro Arg Val Asp Lys Arg
2100 2105 2110
Asn Pro Glu His Cys Val Asn Ser Glu Met Glu Lys Thr Cys Ser Lys
2115 2120 2125
Glu Phe Lys Leu Ser Asn Asn Leu Asn Val Glu Gly Gly Ser Ser Glu
2130 2135 2140
Asn Asn His Ser Ile Lys Val Ser Pro Tyr Leu Ser Gln Phe Gln Gln
2145 2150 2155 2160
Asp Lys Gln Gln Leu Val Leu Gly Thr Lys Val Ser Leu Val Glu Asn
2165 2170 2175
Ile His Val Leu Gly Lys Glu Gln Ala Ser Pro Lys Asn Val Lys Met
2180 2185 2190
Glu Ile Gly Lys Thr Glu Thr Phe Ser Asp Val Pro Val Lys Thr Asn
2195 2200 2205
Ile Glu Val Cys Ser Thr Tyr Ser Lys Asp Ser Glu Asn Tyr Phe Glu
2210 2215 2220
Thr Glu Ala Val Glu Ile Ala Lys Ala Phe Met Glu Asp Asp Glu Leu
2225 2230 2235 2240
Thr Asp Ser Lys Leu Pro Ser His Ala Thr His Ser Leu Phe Thr Cys
2245 2250 2255
Pro Glu Asn Glu Glu Met Val Leu Ser Asn Ser Arg Ile Gly Lys Arg
2260 2265 2270
Arg Gly Glu Pro Leu Ile Leu Val Gly Glu Pro Ser Ile Lys Arg Asn
2275 2280 2285
Leu Leu Asn Glu Phe Asp Arg Ile Ile Glu Asn Gln Glu Lys Ser Leu
2290 2295 2300
Lys Ala Ser Lys Ser Thr Pro Asp Gly Thr Ile Lys Asp Arg Arg Leu
2305 2310 2315 2320
Phe Met His His Val Ser Leu Glu Pro Ile Thr Cys Val Pro Phe Arg
2325 2330 2335
Thr Thr Lys Glu Arg Gln Glu Ile Gln Asn Pro Asn Phe Thr Ala Pro
2340 2345 2350
Gly Gln Glu Phe Leu Ser Lys Ser His Leu Tyr Glu His Leu Thr Leu
2355 2360 2365
Glu Lys Ser Ser Ser Asn Leu Ala Val Ser Gly His Pro Phe Tyr Gln
2370 2375 2380
Val Ser Ala Thr Arg Asn Glu Lys Met Arg His Leu Ile Thr Thr Gly
2385 2390 2395 2400
Arg Pro Thr Lys Val Phe Val Pro Pro Phe Lys Thr Lys Ser His Phe
2405 2410 2415
His Arg Val Glu Gln Cys Val Arg Asn Ile Asn Leu Glu Glu Asn Arg
2420 2425 2430
Gln Lys Gln Asn Ile Asp Gly His Gly Ser Asp Asp Ser Lys Asn Lys
2435 2440 2445
Ile Asn Asp Asn Glu Ile His Gln Phe Asn Lys Asn Asn Ser Asn Gln
2450 2455 2460
Ala Ala Ala Val Thr Phe Thr Lys Cys Glu Glu Pro Leu Asp Leu
2465 2470 2475 2480
Ile Thr Ser Leu Gln Asn Ala Arg Asp Ile Gln Asp Met Arg Ile Lys
2485 2490 2495
Lys Lys Gln Arg Gln Arg Val Phe Pro Gln Pro Gly Ser Leu Tyr Leu
2500 2505 2510
Ala Lys Thr Ser Thr Leu Pro Arg Ile Ser Leu Lys Ala Ala Val Gly
2515 2520 2525
Gly Gln Val Pro Ser Ala Cys Ser His Lys Gln Leu Tyr Thr Tyr Gly
2530 2535 2540
Val Ser Lys His Cys Ile Lys Ile Asn Ser Lys Asn Ala Glu Ser Phe
2545 2550 2555 2560
Gln Phe His Thr Glu Asp Tyr Phe Gly Lys Glu Ser Leu Trp Thr Gly
2565 2570 2575
Lys Gly Ile Gln Leu Ala Asp Gly Gly Trp Leu Ile Pro Ser Asn Asp

2580 2585 2590
 Gly Lys Ala Gly Lys Glu Glu Phe Tyr Arg Ala Leu Cys Asp Thr Pro
 2595 2600 2605
 Gly Val Asp Pro Lys Leu Ile Ser Arg Ile Trp Val Tyr Asn His Tyr
 2610 2615 2620
 Arg Trp Ile Ile Trp Lys Leu Ala Ala Met Glu Cys Ala Phe Pro Lys
 2625 2630 2635 2640
 Glu Phe Ala Asn Arg Cys Leu Ser Pro Glu Arg Val Leu Leu Gln Leu
 2645 2650 2655
 Lys Tyr Arg Tyr Asp Thr Glu Ile Asp Arg Ser Arg Arg Ser Ala Ile
 2660 2665 2670
 Lys Lys Ile Met Glu Arg Asp Asp Thr Ala Ala Lys Thr Leu Val Leu
 2675 2680 2685
 Cys Val Ser Asp Ile Ile Ser Leu Ser Ala Asn Ile Ser Glu Thr Ser
 2690 2695 2700
 Ser Asn Lys Thr Ser Ser Ala Asp Thr Gln Lys Val Ala Ile Ile Glu
 2705 2710 2715 2720
 Leu Thr Asp Gly Trp Tyr Ala Val Lys Ala Gln Leu Asp Pro Pro Leu
 2725 2730 2735
 Leu Ala Val Leu Lys Asn Gly Arg Leu Thr Val Gly Gln Lys Ile Ile
 2740 2745 2750
 Leu His Gly Ala Glu Leu Val Gly Ser Pro Asp Ala Cys Thr Pro Leu
 2755 2760 2765
 Glu Ala Pro Glu Ser Leu Met Leu Lys Ile Ser Ala Asn Ser Thr Arg
 2770 2775 2780
 Pro Ala Arg Trp Tyr Thr Lys Leu Gly Phe Phe Pro Asp Pro Arg Pro
 2785 2790 2795 2800
 Phe Pro Leu Pro Leu Ser Ser Leu Phe Ser Asp Gly Gly Asn Val Gly
 2805 2810 2815
 Cys Val Asp Val Ile Ile Gln Arg Ala Tyr Pro Ile Gln Trp Met Glu
 2820 2825 2830
 Lys Thr Ser Ser Gly Leu Tyr Ile Phe Arg Asn Glu Arg Glu Glu Glu
 2835 2840 2845
 Lys Glu Ala Ala Lys Tyr Val Glu Ala Gln Gln Lys Arg Leu Glu Ala
 2850 2855 2860
 Leu Phe Thr Lys Ile Gln Glu Glu Phe Glu Glu His Glu Glu Asn Thr
 2865 2870 2875 2880
 Thr Lys Pro Tyr Leu Pro Ser Arg Ala Leu Thr Arg Gln Gln Val Arg
 2885 2890 2895
 Ala Leu Gln Asp Gly Ala Glu Leu Tyr Glu Ala Val Lys Asn Ala Ala
 2900 2905 2910
 Asp Pro Ala Tyr Leu Glu Gly Tyr Phe Ser Glu Glu Gln Leu Arg Ala
 2915 2920 2925
 Leu Asn Asn His Arg Gln Met Leu Asn Asp Lys Lys Gln Ala Gln Ile
 2930 2935 2940
 Gln Leu Glu Ile Arg Lys Thr Met Glu Ser Ala Glu Gln Lys Glu Gln
 2945 2950 2955 2960
 Gly Leu Ser Arg Asp Val Thr Thr Val Trp Lys Leu Arg Ile Val Ser
 2965 2970 2975
 Tyr Ser Lys Lys Glu Lys Asp Ser Val Ile Leu Ser Ile Trp Arg Pro
 2980 2985 2990
 Ser Ser Asp Leu Tyr Ser Leu Leu Thr Glu Gly Lys Arg Tyr Arg Ile
 2995 3000 3005
 Tyr His Leu Ala Thr Ser Lys Ser Lys Ser Lys Ser Glu Arg Ala Asn
 3010 3015 3020
 Ile Gln Leu Ala Ala Thr Lys Lys Thr Gln Tyr Gln Gln Leu Pro Val
 3025 3030 3035 3040
 Ser Asp Glu Ile Leu Phe Gln Ile Tyr Gln Pro Arg Glu Pro Leu His
 3045 3050 3055
 Phe Ser Lys Phe Leu Asp Pro Asp Phe Gln Pro Ser Cys Ser Glu Val
 3060 3065 3070

D
CMTK

Asp Leu Ile Gly Phe Val Val Ser Val Val Lys Lys Thr Gly Leu Ala
3075 3080 3085
Pro Phe Val Tyr Leu Ser Asp Glu Cys Tyr Asn Leu Leu Ala Ile Lys
3090 3095 3100
Phe Trp Ile Asp Leu Asn Glu Asp Ile Ile Lys Pro His Met Leu Ile
3105 3110 3115 3120
Ala Ala Ser Asn Leu Gln Trp Arg Pro Glu Ser Lys Ser Gly Leu Leu
3125 3130 3135
Thr Leu Phe Ala Gly Asp Phe Ser Val Phe Ser Ala Ser Pro Lys Glu
3140 3145 3150
Gly His Phe Gln Glu Thr Phe Asn Lys Met Lys Asn Thr Val Glu Asn
3155 3160 3165
Ile Asp Ile Leu Cys Asn Glu Ala Glu Asn Lys Leu Met His Ile Leu
3170 3175 3180
His Ala Asn Asp Pro Lys Trp Ser Thr Pro Thr Lys Asp Cys Thr Ser
3185 3190 3195 3200
Gly Pro Tyr Thr Ala Gln Ile Ile Pro Gly Thr Gly Asn Lys Leu Leu
3205 3210 3215
Met Ser Ser Pro Asn Cys Glu Ile Tyr Tyr Gln Ser Pro Leu Ser Leu
3220 3225 3230
Cys Met Ala Lys Arg Lys Ser Val Ser Thr Pro Val Ser Ala Gln Met
3235 3240 3245
Thr Ser Lys Ser Cys Lys Gly Glu Lys Glu Ile Asp Asp Gln Lys Asn
3250 3255 3260
Cys Lys Lys Arg Arg Ala Leu Asp Phe Leu Ser Arg Leu Pro Leu Pro
3265 3270 3275 3280
Pro Pro Val Ser Pro Ile Cys Thr Phe Val Ser Pro Ala Ala Gln Lys
3285 3290 3295
Ala Phe Gln Pro Pro Arg Ser Cys Gly Thr Lys Tyr Glu Thr Pro Ile
3300 3305 3310
Lys Lys Glu Leu Asn Ser Pro Gln Met Thr Pro Phe Lys Lys Phe
3315 3320 3325
Asn Glu Ile Ser Leu Leu Glu Ser Asn Ser Ile Ala Asp Glu Glu Leu
3330 3335 3340
Ala Leu Ile Asn Thr Gln Ala Leu Leu Ser Gly Ser Thr Gly Glu Lys
3345 3350 3355 3360
Gln Phe Ile Ser Val Ser Glu Ser Thr Arg Thr Ala Pro Thr Ser Ser
3365 3370 3375
Glu Asp Tyr Leu Arg Leu Lys Arg Arg Cys Thr Thr Ser Leu Ile Lys
3380 3385 3390
Glu Gln Glu Ser Ser Gln Ala Ser Thr Glu Glu Cys Glu Lys Asn Lys
3395 3400 3405
Gln Asp Thr Ile Thr Thr Lys Lys Tyr Ile
3410 3415

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE:

- (A) NAME/KEY: other
- (B) LOCATION: 1...20
- (D) OTHER INFORMATION: 2F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

TGAGTTTAC CTCAGTCACA

20

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 41 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

CAGGAAACAG CTATGACCCT GTGACGTACT GGGTTTTAG C

41

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

D
DK
CMK
(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: other
- (B) LOCATION: 1...20
- (D) OTHER INFORMATION: 3FII primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

GATCTTTAAC TGTTCTGGGT CACA

24

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: other
- (B) LOCATION: 1...22
- (D) OTHER INFORMATION: 3RII primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

CCCAGCATGA CACAATTAAAT GA

22

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 44 base pairs
- (B) TYPE: nucleic acid

- (C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

- (ix) FEATURE
(A) NAME/KEY: other
(B) LOCATION: 1...44
(D) OTHER INFORMATION: 4F/M 13F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

TGTAAACGA CGGCCAGTAG AATGCAAATT TATAATCCAG AGTA

44

(2) INFORMATION FOR SEQ ID NO:19:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

D
cont
(ii) MOLECULE TYPE: Genomic DNA

- (ix) FEATURE
(A) NAME/KEY: other
(B) LOCATION: 1...22
(D) OTHER INFORMATION: 4R-1A primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

ATCAGATTCA TCTTTATAGA AC

22

(2) INFORMATION FOR SEQ ID NO:20:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 40 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

- (ix) FEATURE
(A) NAME/KEY: other
(B) LOCATION: 1...40
(D) OTHER INFORMATION: 5+6F/M13F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

TGTAAACGA CGGCCAGTTG TGTTGGCATT TTAAACATCA

40

(2) INFORMATION FOR SEQ ID NO:21:

- (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 38 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

- (ix) FEATURE
 - (A) NAME/KEY: other
 - (B) LOCATION: 1...38
 - (D) OTHER INFORMATION: 5+6R/M13R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

CAGGAAACAG CTATGACCCA GGGCAAAGGT ATAACGCT

38

(2) INFORMATION FOR SEQ ID NO:22:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 38 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

D
W
(ii) MOLECULE TYPE: Genomic DNA

- (ix) FEATURE
 - (A) NAME/KEY: other
 - (B) LOCATION: 1...38
 - (D) OTHER INFORMATION: 7F/M13F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

TGTAAAACGA CGGCCAGTTA AGTGAAATAA AGAGTGAA

38

(2) INFORMATION FOR SEQ ID NO:23:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 36 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

- (ix) FEATURE
 - (A) NAME/KEY: other
 - (B) LOCATION: 1...36
 - (D) OTHER INFORMATION: 7R/M13R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

CAGGAAACAG CTATGACCAAG AAGTATTAGA GATGAC

36

(2) INFORMATION FOR SEQ ID NO:24:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 40 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

- (ix) FEATURE
- (A) NAME/KEY: other
 - (B) LOCATION: 1...40
 - (D) OTHER INFORMATION: 8F/M13F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

TGTAAAACGA CGGCCAGTGC CATATCTTAC CACCTTGTGA

40

D
J
CMK
(2) INFORMATION FOR SEQ ID NO:25:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 22 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: other
- (B) LOCATION: 1...22
- (D) OTHER INFORMATION: 8FIA primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

TTGCATTCTA GTGATAATAT AC

22

(2) INFORMATION FOR SEQ ID NO:26:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 19 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

- (ix) FEATURE
- (A) NAME/KEY: other
 - (B) LOCATION: 1...19
 - (D) OTHER INFORMATION: 8RIA primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

AATTGTTAGC AATTCAAC

19

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 40 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: other
- (B) LOCATION: 1...40
- (D) OTHER INFORMATION: 9F/M13F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

TGTAAAACGA CGGCCAGTTG GACCTAGGTT GATTGCAGAT

40

D
1
CONT
(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 40 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: other
- (B) LOCATION: 1...40
- (D) OTHER INFORMATION: 9R/M13R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

CAGGAAACAG CTATGACCTA AACTGAGATC ACAGGGTGACA

40

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: other
- (B) LOCATION: 1...24
- (D) OTHER INFORMATION: 10AF primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

GAATAATATA AATTATATGG CTTA

24

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 37 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: other
- (B) LOCATION: 1...37
- (D) OTHER INFORMATION: 10AR/M13R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

CAGGAAACAG CTATGACCCC TAGTCTTGCT AGTTCTT

37

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 42 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: other
- (B) LOCATION: 1...42
- (D) OTHER INFORMATION: 10BF/M13F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

TGTAAAACGA CGGCCAGTAR CTGAAGTGGA ACCAAATGAT AC

42

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 44 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: other
- (B) LOCATION: 1...44
- (D) OTHER INFORMATION: 10BR/M13R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

CAGGAAACAG CTATGACCAC GTGGCAAAGA ATTCTCTGAA GTAA

44

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 40 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ix) FEATURE:

(ii) MOLECULE TYPE: Genomic DNA

- (A) NAME/KEY: other
- (B) LOCATION: 1...40
- (D) OTHER INFORMATION: 10CF/M13F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

TGTAAAACGA CGGCCAGTCA GCATCTTGAA TCTCATACAG

40

D
Cont
(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: other
- (B) LOCATION: 1...19
- (D) OTHER INFORMATION: 10CRII primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

AGACAGAGGT ACCTGAATC

19

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 40 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: other
- (B) LOCATION: 1...40
- (D) OTHER INFORMATION: 11AF-M13 primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

TGTAAAACGA CGGCCAGTTG GTACTTTAAT TTTGTCACCTT

40

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 37 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: other
- (B) LOCATION: 1...37
- (D) OTHER INFORMATION: 11AR-M13 primer

D
CK
CPK
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

CAGGAAACAG CTATGACCTG CAGGCATGAC AGAGAAT

37

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: other
- (B) LOCATION: 1...22
- (D) OTHER INFORMATION: 11BF primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

AAGAAGCAAA ATGTAATAAG GA

22

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: other
- (B) LOCATION: 1...22

(D) OTHER INFORMATION: 11BR primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

CATTTAAAGC ACATAACATCT TG

22

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: primer
- (B) LOCATION:
- (D) OTHER INFORMATION: 11CF primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

TCTAGAGGCA AAGAACATA C

21

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: other
- (B) LOCATION: 1...22
- (D) OTHER INFORMATION: 11CR primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

CAAGATTATT CCTTTCATTA GC

22

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: other
(B) LOCATION: 1...22
(D) OTHER INFORMATION: 11DF primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

AACCAAAACA CAAATCTAAG AG

22

(2) INFORMATION FOR SEQ ID NO:42:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: other
(B) LOCATION: 1...23
(D) OTHER INFORMATION: 11DR primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

GTCATTTTTA TATGCTGCTT TAC

23

(2) INFORMATION FOR SEQ ID NO:43:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: other
(B) LOCATION: 1...21
(D) OTHER INFORMATION: 11EF primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

GGTTTATAT GGAGACACAG G

21

(2) INFORMATION FOR SEQ ID NO:44:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

- (ix) FEATURE
(A) NAME/KEY: other
(B) LOCATION: 1...23
(D) OTHER INFORMATION: 11ER primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

GTATTTACAA TTTCAACACA AGC

23

(2) INFORMATION FOR SEQ ID NO:45:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

- D
CONK
(ix) FEATURE
(A) NAME/KEY: other
(B) LOCATION: 1...20
(D) OTHER INFORMATION: 11FF primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

ATCACAGTTT TGGAGGTAGC

20

(2) INFORMATION FOR SEQ ID NO:46:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

- (ix) FEATURE
(A) NAME/KEY: other
(B) LOCATION: 1...21
(D) OTHER INFORMATION: 11FR primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

CTGACTTCCT GATTCTTCTA A

21

(2) INFORMATION FOR SEQ ID NO:47:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: other
- (B) LOCATION: 1...21
- (D) OTHER INFORMATION: 11GF primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

CTCAGATGTT ATTTTCCAAG C

21

(2) INFORMATION FOR SEQ ID NO:48:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: other
- (B) LOCATION: 1...21
- (D) OTHER INFORMATION: 11GR primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

CTGTTAAATA ACCAGAAGCA C

21

(2) INFORMATION FOR SEQ ID NO:49:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: other
- (B) LOCATION: 1...18
- (D) OTHER INFORMATION: 11HF primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

AGGTAGACAG CAGCAAGC

18

(2) INFORMATION FOR SEQ ID NO:50:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ix) FEATURE:

- (A) NAME/KEY: other
- (B) LOCATION: 1...22
- (D) OTHER INFORMATION: 11HR primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

GTAATATCAG TTGGCATTAA TT

22

(2) INFORMATION FOR SEQ ID NO:51:

- D
CONK
- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 21 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

- (A) NAME/KEY: other
- (B) LOCATION: 1...21
- (D) OTHER INFORMATION: 11IF primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

TGCAGAGGTA CATCCAATAA G

21

(2) INFORMATION FOR SEQ ID NO:52:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: other
- (B) LOCATION: 1...21
- (D) OTHER INFORMATION: 11IR primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

GATCAGTAAA TAGCAAGTCC G

21

(2) INFORMATION FOR SEQ ID NO:53:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: other
- (B) LOCATION: 1...23
- (D) OTHER INFORMATION: 11JF primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

TACTGAAAAT GAAGATAACA AAT

23

(2) INFORMATION FOR SEQ ID NO:54:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

D
1
X
cont
(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: other
- (B) LOCATION: 1...22
- (D) OTHER INFORMATION: 11JR primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

ATTTTGTTCT TTCTTATGTC AG

22

(2) INFORMATION FOR SEQ ID NO:55:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: other
- (B) LOCATION: 1...35
- (D) OTHER INFORMATION: 11KF-M13 primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

TGTAAACCGA CGGCCAGTCT ACTAAAACGG AGCAA

35

(2) INFORMATION FOR SEQ ID NO:56:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

- (A) NAME/KEY: other
(B) LOCATION: 1...35
(D) OTHER INFORMATION: 11KR-M13 primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

CAGGAAACAG CTATGACCGT ATGAAAACCC AACAG

35

(2) INFORMATION FOR SEQ ID NO:57:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

- (ix) FEATURE
(A) NAME/KEY: other
(B) LOCATION: 1...22
(D) OTHER INFORMATION: 11LF primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

CACAAAATAAC TGAAAGAAAG TG

22

(2) INFORMATION FOR SEQ ID NO:58:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

- (ix) FEATURE
(A) NAME/KEY: other
(B) LOCATION: 1...19
(D) OTHER INFORMATION: 11LR primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:

GGCACCAACAG TCTCAATAG

19

(2) INFORMATION FOR SEQ ID NO:59:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: other
- (B) LOCATION: 1...20
- (D) OTHER INFORMATION: 11MF primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:59:

GCAAAGACCC TAAAGTACAG

20

(2) INFORMATION FOR SEQ ID NO:60:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: other
- (B) LOCATION: 1...22
- (D) OTHER INFORMATION: 11MR primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:60:

CATCAAATAT TCCTTCTCTA AG

22

(2) INFORMATION FOR SEQ ID NO:61:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: other
- (B) LOCATION: 1...35
- (D) OTHER INFORMATION: 11NF-M13 primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:61:

TGTAAAACGA CGGCCAGTGA AAATTCAAGCC TTAGC

35

(2) INFORMATION FOR SEQ ID NO:62:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: other
- (B) LOCATION: 1...35
- (D) OTHER INFORMATION: 11NR-M13 primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:62:

CAGGAAACAG CTATGACCAT CAGAATGGTA GGAAT

35

(2) INFORMATION FOR SEQ ID NO:63:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

D
CONK
(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: other
- (B) LOCATION: 1...22
- (D) OTHER INFORMATION: 11OF primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:63:

GTACTATAGC TGAAAATGAC AA

22

(2) INFORMATION FOR SEQ ID NO:64:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: other
- (B) LOCATION: 1...20
- (D) OTHER INFORMATION: 11OR primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:64:

ACCACTGGCT ATCCTAAATG

20

(2) INFORMATION FOR SEQ ID NO:65:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs

- (B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

- (ix) FEATURE
(A) NAME/KEY: other
(B) LOCATION: 1...20
(D) OTHER INFORMATION: 11PF primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:65:

TGAAGATATT TGCAGTTGAGG

20

(2) INFORMATION FOR SEQ ID NO:66:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

- (ix) FEATURE
(A) NAME/KEY: other
(B) LOCATION: 1...20
(D) OTHER INFORMATION: 11PR primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:66:

GTCAGCAAAA ACCTTATGTG

20

(2) INFORMATION FOR SEQ ID NO:67:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

- (ix) FEATURE
(A) NAME/KEY: other
(B) LOCATION: 1...21
(D) OTHER INFORMATION: 11QF primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:67:

ACGAAAATTA TGGCAGGTTG T

21

(2) INFORMATION FOR SEQ ID NO:68:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

- (ix) FEATURE
(A) NAME/KEY: other
(B) LOCATION: 1...21
(D) OTHER INFORMATION: 11QR primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:68:

CTTGTCTTGC GTTTGTAAT G

21

(2) INFORMATION FOR SEQ ID NO:69:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

- (ix) FEATURE
(A) NAME/KEY: other
(B) LOCATION: 1...20
(D) OTHER INFORMATION: 11RF primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:69:

GCTTCATAAG TCAGTCTCAT

20

(2) INFORMATION FOR SEQ ID NO:70:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

- (ix) FEATURE
(A) NAME/KEY: other
(B) LOCATION: 1...20
(D) OTHER INFORMATION: 11RR primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:70:

TCAAATTCCCT CTAACACTCC

20

(2) INFORMATION FOR SEQ ID NO:71:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 35 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: other
- (B) LOCATION: 1...35
- (D) OTHER INFORMATION: 11SF-M13 primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:71:

TGTAAAACGA CGGCCAGTTA CAGCAAGTGG AAAGC

35

D
Cont
(2) INFORMATION FOR SEQ ID NO:72:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 37 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: other
- (B) LOCATION: 1...37
- (D) OTHER INFORMATION: 11SR-M13 primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:72:

CAGGAAACAG CTATGACCAA GTTTCAGTTT TACCAAT

37

(2) INFORMATION FOR SEQ ID NO:73:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 22 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: other
- (B) LOCATION: 1...22
- (D) OTHER INFORMATION: 11TF primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:73:

GTTCTTCAGA AAATAATCAC TC

22

(2) INFORMATION FOR SEQ ID NO:74:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: other
- (B) LOCATION: 1...21
- (D) OTHER INFORMATION: 11TR primer

D
1
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:74:

TGTAAAAAGA GAATGTGTGG C

21

(2) INFORMATION FOR SEQ ID NO:75:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: other
- (B) LOCATION: 1...39
- (D) OTHER INFORMATION: 11UF-M13 primer

D
Cmt
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:75:

TGTAAAACGA CGGCCAGTAC TTTTCTGAT GTTCCTGTG

39

(2) INFORMATION FOR SEQ ID NO:76:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: other
- (B) LOCATION: 1...39
- (D) OTHER INFORMATION: 11UR-M13 primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:76:

CAGGAAACAG CTATGACCTA AAAATAGTGA TTGGCAACA

39

(2) INFORMATION FOR SEQ ID NO:77:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 42 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: other
- (B) LOCATION: 1...42
- (D) OTHER INFORMATION: 12F/M13F primer

D
R
J
N
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:77:

TGTAAAACGA CGGCCAGTAG TGGTGTTTA AAGTGGTCAA AA

42

(2) INFORMATION FOR SEQ ID NO:78:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 40 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: other
- (B) LOCATION: 1...40
- (D) OTHER INFORMATION: 12R/M13R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:78:

CAGGAAACAG CTATGACCGG ATCCACCTGA GGTCAGAATA

40

(2) INFORMATION FOR SEQ ID NO:79:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: other

- (B) LOCATION: 1...21
(D) OTHER INFORMATION: 13-2F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:79:

TAACATTTAA GCATCCGTTA C

21

(2) INFORMATION FOR SEQ ID NO:80:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

- D
Cont
(ix) FEATURE
(A) NAME/KEY: other
(B) LOCATION: 1...28
(D) OTHER INFORMATION: 13-2R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:80:

AAACGAGACT TTTCTCATAC TGTATTAG

28

(2) INFORMATION FOR SEQ ID NO:81:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

- (ix) FEATURE
(A) NAME/KEY: other
(B) LOCATION: 1...22
(D) OTHER INFORMATION: 14F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:81:

ACCATGTAGC AAATGAGGGT CT

22

(2) INFORMATION FOR SEQ ID NO:82:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

- (ix) FEATURE
(A) NAME/KEY: other
(B) LOCATION: 1...22
(D) OTHER INFORMATION: 14AR primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:82:

GCTTTTGTCT GTTTTCCTCC AA

22

(2) INFORMATION FOR SEQ ID NO:83:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

- D
D
CONK
(ix) FEATURE
(A) NAME/KEY: other
(B) LOCATION: 1...21
(D) OTHER INFORMATION: 15-2F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:83:

CCAGGGGTTG TGCTTTTTAA A

21

(2) INFORMATION FOR SEQ ID NO:84:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 38 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

- (ix) FEATURE
(A) NAME/KEY: primer
(B) LOCATION:
(D) OTHER INFORMATION: 15FUT/M13-R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:84:

CAGGAAACAG CTATGACCAC TCTGTCATAA AAGCCATC

38

(2) INFORMATION FOR SEQ ID NO:85:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: other
- (B) LOCATION: 1...24
- (D) OTHER INFORMATION: 16AF primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:85:

TTGGGTTTGT TATAATTGTT TTTA

24

(2) INFORMATION FOR SEQ ID NO:86:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

D
ix
COMX
(ix) FEATURE

- (A) NAME/KEY: other
- (B) LOCATION: 1...20
- (D) OTHER INFORMATION: 16AR primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:86:

CCAACTTTT AGTCGAGAG

20

(2) INFORMATION FOR SEQ ID NO:87:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

D
ix
COMX
(ix) FEATURE

- (A) NAME/KEY: other
- (B) LOCATION: 1...19
- (D) OTHER INFORMATION: 17F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:87:

TTCAGTATCA TCCTATGTG

19

(2) INFORMATION FOR SEQ ID NO:88:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: other
- (B) LOCATION: 1...20
- (D) OTHER INFORMATION: 17AR primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:88:

AGAAACCTTA ACCCATACTG

20

(2) INFORMATION FOR SEQ ID NO:89:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

D
Cont
(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: other
- (B) LOCATION: 1...39
- (D) OTHER INFORMATION: 18FUT/M13-AF primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:89:

TGTAAAACGA CGGCCAGTGA ATTCTAGAGT CACACTTCC

39

(2) INFORMATION FOR SEQ ID NO:90:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 38 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: other
- (B) LOCATION: 1...38
- (D) OTHER INFORMATION: 18R/M13R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:90:

CAGGAAACAG CTATGACCTT TAACTGAATC AATGACTG

38

(2) INFORMATION FOR SEQ ID NO:91:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 41 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: other
- (B) LOCATION: 1...41
- (D) OTHER INFORMATION: 19F/M13F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:91:

TGTAAAACGA CGGCCAGTAA GTGAATATT TTAAGGCAGT T

41

(2) INFORMATION FOR SEQ ID NO:92:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: other
- (B) LOCATION: 1...39
- (D) OTHER INFORMATION: 19FUT/M13-R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:92:

CAGGAAACAG CTATGACCAA GAGACCGAAA CTCCATCTC

39

(2) INFORMATION FOR SEQ ID NO:93:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 38 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: other
- (B) LOCATION: 1...38
- (D) OTHER INFORMATION: 20F/M13F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:93:

TGTAAAACGA CGGCCAGTCA CTGTGCCTGG CCTGATAC

38

(2) INFORMATION FOR SEQ ID NO:94:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39 base pairs

- (B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

- (ix) FEATURE
(A) NAME/KEY: other
(B) LOCATION: 1...39
(D) OTHER INFORMATION: 20R/M13R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:94:

CAGGAAACAG CTATGACCAT GTTAAATTCA AAGTCTCTA

39

(2) INFORMATION FOR SEQ ID NO:95:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 39 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

D work
(ii) MOLECULE TYPE: Genomic DNA

- (ix) FEATURE
(A) NAME/KEY: other
(B) LOCATION: 1...39
(D) OTHER INFORMATION: 21F/M13F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:95:

TGTAAAACGA CGGCCAGTGG GTGTTTATG CTTGGTTCT

39

(2) INFORMATION FOR SEQ ID NO:96:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 40 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

- (ix) FEATURE
(A) NAME/KEY: other
(B) LOCATION: 1...40
(D) OTHER INFORMATION: 21R/M13R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:96:

CAGGAAACAG CTATGACCCA TTTAACACATA TTCCTTCCTG

40

(2) INFORMATION FOR SEQ ID NO:97:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: other
- (B) LOCATION: 1...19
- (D) OTHER INFORMATION: 22F-1A primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:97:

AACCACACCC TTAAGATGA

19

D
✓
✓

(2) INFORMATION FOR SEQ ID NO:98:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: other
- (B) LOCATION: 1...20
- (D) OTHER INFORMATION: 22R-1A primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:98:

GCATTAGTAG TGGATTTGCG

20

(2) INFORMATION FOR SEQ ID NO:99:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 16 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: other
- (B) LOCATION: 1...16
- (D) OTHER INFORMATION: 23FII primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:99:

TCACCTCCAT TGCATC

16

(2) INFORMATION FOR SEQ ID NO:100:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: other
- (B) LOCATION: 1...17
- (D) OTHER INFORMATION: 23RII primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:100:

TGCCAACTGG TAGCTCC

17

D
Don't
Count
(2) INFORMATION FOR SEQ ID NO:101:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: other
- (B) LOCATION: 1...20
- (D) OTHER INFORMATION: 24 2F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:101:

TACAGTTAGC AGCGACAAAAA

20

(2) INFORMATION FOR SEQ ID NO:102:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 38 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: other
- (B) LOCATION: 1...38
- (D) OTHER INFORMATION: 24R/M13R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:102:

CAGGAAACAG CTATGACCAT TTGCCAACTG GTAGCTCC

38

(2) INFORMATION FOR SEQ ID NO:103:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: other
- (B) LOCATION: 1...20
- (D) OTHER INFORMATION: 25F-7/23 primer

D
D
COMT
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:103:

GCTTTCGCCA AATTCAGCTA

20

(2) INFORMATION FOR SEQ ID NO:104:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: other
- (B) LOCATION: 1...20
- (D) OTHER INFORMATION: 25R-7/23 primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:104:

TACCAAAATG TGTGGTGATG

20

(2) INFORMATION FOR SEQ ID NO:105:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: other

- (B) LOCATION: 1...20
(D) OTHER INFORMATION: 26-2F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:105:

AATCACTGAT ACTGGTTTG

20

(2) INFORMATION FOR SEQ ID NO:106:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

- (ix) FEATURE
(A) NAME/KEY: other
(B) LOCATION: 1...20
(D) OTHER INFORMATION: 26-2R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:106:

TATACTTACA GGAGCCACAT

20

(2) INFORMATION FOR SEQ ID NO:107:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

- (ix) FEATURE
(A) NAME/KEY: other
(B) LOCATION: 1...18
(D) OTHER INFORMATION: 27AF-1A primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:107:

CTGTGTGTA TATTTGCG

18

(2) INFORMATION FOR SEQ ID NO:108:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 40 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: other
(B) LOCATION: 1...40
(D) OTHER INFORMATION: 27AR/M13R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:108:

CAGGAAACAG CTATGACGGC AAGTTCTTCG TCAGCTATTG

40

(2) INFORMATION FOR SEQ ID NO:109:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 40 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

D
|
D
Cont
(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: other
(B) LOCATION: 1...40
(D) OTHER INFORMATION: 27BF/M13F primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:109:

TGTAAAACGA CGGCCAGTGA ATTCTCCTCA GATGACTCCA

40

(2) INFORMATION FOR SEQ ID NO:110:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 38 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: other
(B) LOCATION: 1...38
(D) OTHER INFORMATION: 27BR/M13R primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:110:

CAGGAAACAG CTATGACCTC TTTGCTCATT GTGCAACA

38

(2) INFORMATION FOR SEQ ID NO:111:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1158 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

(A) NAME/KEY: exon

(B) LOCATION: 1...1158

(D) OTHER INFORMATION: Exon 10, nucleotides 31-1146 correspond to nucleotides 1022-2137 of complete coding sequence

(ix) FEATURE

(A) NAME/KEY: variable

(B) LOCATION: 102...602

(D) OTHER INFORMATION: M at positions 102 or 351 = A or C; R at position 602 = A or G.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:111:

D'cont

TTAATGTGCT TCTGTTTAT ACTTTAACAG GATTTGGAAA AACATCAGGG AATTCAATT 60
AAGTAAATAG CTGCAAAGAC CACATTGGAA AGTCAATGCC AMATGTCCTA GAAGATGAAG 120
TATATGAAAC AGTTGTAGAT ACCTCTGAAG AAGATAGTTT TTCATTATGT TTTTCTAAAT 180
GTAGAACAAA AAATCTACAA AAAGTAAGAA CTAGCAAGAC TAGGAAAAAA ATTTTCCATG 240
AAGCAAACGC TGATGAATGT GAAAAATCTA AAAACCAAGT GAAAGAAAAAA TACTCATTG 300
TATCTGAAGT GGAACCAAAT GATACTGATC CATTAGATT AAATGTAGCA MATCAGAAC 360
CCTTGAGAG TGGAAGTGAC AAAATCTCCA AGGAAGTTGT ACCGTCTTG GCCTGTGAAT 420
GGTCTCAACT AACCCCTTC A GGTCTAAATG GAGCCCAGAT GGAGAAAATA CCCCTATTG 480
ATATTTCTTC ATGTGACCAA AATATTTCAG AAAAAGACCT ATTAGACACA GAGAACAAAA 540
GAAAGAAAGA TTTTCTTACT TCAGAGAATT CTTGCCACG TATTCTTAGC CTACCAAAAT 600
CRGAGAAGCC ATAAAATGAG GAAACAGTGG TAAATAAGAG AGATGAAGAG CAGCATCTG 660
AATCTCATAC AGACTGCATT CTTGCAGTAA AGCAGGCAAT ATCTGGAAC TCTCCAGTGG 720
CTTCTTCATT TCAGGGTATC AAAAAGCTA TATTCTAGAAT AAGAGAATCA CCTAAAGAGA 780
CTTCAATGC AAGTTTTCA GGTCTATGA CTGATCCAAA CTTTAAAAAA GAAACTGAAG 840
CCTCTGAAAG TGGACTGGAA ATACATACTG TTTGCTCACA GAAGGAGGAC TCCTTATGTC 900
CAAATTTAAT TGATAATGGA AGCTGGCCAG CCACCACAC ACAGAATTCT GTAGCTTTGA 960
AGAATGCAGG TTAAATATCC ACTTTGAAAAA AGAAAACAAA TAAGTTTATT TATGCTATAC 1020
ATGATGAAAC ATTTTATAAA GGAAAAAAA TACCGAAAGA CCAAAATCA GAACTAATTA 1080
ACTGTTCAGC CCAGTTGAA GCAAATGCTT TTGAAGCACC ACTTACATT GCAAATGCTG 1140
ATTCAAGGTAC CTCTGTCT 1158

(2) INFORMATION FOR SEQ ID NO:112:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 4987 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

(A) NAME/KEY: exon

(B) LOCATION: 1...4987

(D) OTHER INFORMATION: Exon 11, nucleotides 20-4951 correspond to nucleotides 2138-7069 of complete coding sequence

(ix) FEATURE

(A) NAME/KEY: variable

(B) LOCATION: 339...1917

(D) OTHER INFORMATION: Y at positions 339 or 1917 = C or T; R at positions 790, 1081 or 1506 = A or G.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:112:

TTTGTGTTT TATGTTAGG TTTATTGCAT TCTTCTGTGA AAAGAAGCTG TTCACAGAAT 60
 GATTCTGAAG AACCAACTTT GTCCTTAAC AGCTCTTTG GGACAATTCT GAGGAAATGT 120
 TCTAGAAATG AAACATGTT TAATAATACA GTAATCTCTC AGGATCTTGA TTATAAAAGAA 180
 GCAAAATGTA ATAAGGAAAA ACTACAGTTA TTTATTACCC CAGAAGCTGA TTCTCTGTCA 240
 TGCCCTGCAGG AAGGACAGTG TGAAAATGAT CCAAAAAGCA AAAAAGTTTC AGATATAAAA 300
 GAAGAGGTCT TGGCTGCAGC ATGTCACCCA GTACAACAYT CAAAAGTGGA ATACAGTGAT 360
 ACTGACTTTC AATCCCAGAA AAGTCTTTA TATGATCATG AAAATGCCAG CACTCTTATT 420
 TTAACTCCTA CTTCCAAGGA TGTTCTGTCA AACCTAGTCA TGATTCTAG AGGCAAAGAA 480
 TCATACAAAAA TGTCAGACAA GCTCAAAGGT ACAATTATG AATCTGATGT TGAATTAACC 540
 AAAAATATTC CCATGGAAAA GAATCAAGAT GTATGTGCTT TAAATGAAAA TTATAAAAC 600
 GTTGAGCTGT TGCCACCTGA AAAATACATG AGAGTAGCAT CACCTTCAAG AAAGGTACAA 660
 TTCAACCAAA ACACAAATCT AAGAGTAATC CAAAAAAATC AAGAAGAAC TACTTCAATT 720
 TCAAAAATAA CTGTCATCC AGACTCTGAA GAACCTTTCT CAGACAATGA GAATAATTTC 780
 GTCTTCAAR TAGCTAATGA AAGGAATAAT CTTGCTTTAG GAAATACTAA GGAACCTTCAT 840
 GAAACAGACT TGACTTGTGT AAACGAACCC ATTTTCAAGA ACTCTACCAT GGTTTTATAT 900
 GGAGACACAG GTGATAAACAG AGCAACCCAA GTGTCAATT AAAAAGATTG GGTTTATGTT 960
 CTTGCAGAGG AGAACAAAAA TAGTGTAAAG CACCATATAA AAATGACTCT AGGTCAAGAT 1020
 TTAAAATCGG ACATCTCCTT GAATATAGAT AAAATACCAAG AAAAATATAA TGATTACATG 1080
 RACAAATGGG CAGGACTCTT AGGTCCAATT TCAAATCACA GTTTGGAGG TAGCTTCAGA 1140
 ACAGCTTCAA ATAAGGAAAT CAAGCTCTCT GAACATAACA TTAAGAAGAG CAAATGTTC 1200
 TTCAAAAGATA TTGAAGAACAA ATATCCTACT AGTTTAGCTT GTGTGAAAT TGTAATTAACC 1260
 TTGGCATTAG ATAATCAAAA GAAACTGAGC AAGCCTCAGT CAATTAAATAC TGTATCTGCA 1320
 CATTACAGA GTAGTGTAGT TGTTCTGAT TGAAAATATA GTCATATAAC CCCTCAGATG 1380
 TTATTTCCA AGCAGGATT TAATTCAAAC CATAATTAA CACCTAGCCA AAAGGCAGAA 1440
 ATTACAGAAC TTTCTACTAT ATTAGAAGAA TCAGGAAGTC AGTTTGAAATT TACTCAGTTT 1500
 AGAAARCCAA GCTACATATT GCAGAAGAGT ACATTTGAAG TGCCTGAAAAA CCAGATGACT 1560
 ATCTTAAAGA CCACTTCTGA GGAATGCAGA GATGCTGATC TTCATGTCAT AATGAATGCC 1620
 CCATCGATTG GTCAGGTAGA CAGCAGCAAG CAATTTGAAG GTACAGTTGA AATTAAACGG 1680
 AAGTTTGCCTG GCCTGTTGAA AAATGACTGT AACAAAAGTG CTTCTGGTTA TTTAACAGAT 1740
 GAAAATGAAG TGGGGTTTAG GGGCTTTAT TCTGCTCATG GCACAAAACT GAATGTTCT 1800
 ACTGAAGCTC TGCAAAAAGC TGTGAAACTG TTTAGTGATA TTGAGAATAT TAGTGAGGAA 1860
 ACTTCTGCAG AGGTACATCC AATAAGTTA TCTTCAAGTA AATGTATGAA TTCTGTYGTT 1920
 TCAATGTTA AGATAGAAAA TCATAATGAT AAAACTGTAA GTGAAAAAAA TAATAATGCA 1980
 CAACTGATAT TACAAAATAA TATTGAAATG ACTACTGGCA CTTTGTGTTA AGAAATTACT 2040
 GAAAATTACA AGAGAAATAC TGAAAATGAA GATAACAAAT ATACTGCTGC CAGTAGAAAT 2100
 TCTCATAACT TAGAATTGAA TGGCAGTGAT TCAAGTAAAA ATGATACTGT TTGTATTCT 2160
 AAAGATGAAA CGGACTTGCAT ATTTACTGAT CAGCACAACA TATGTCTTAA ATTATCTGGC 2220
 CAGTTTATGA AGGAGGGAAA CACTCAGATT AAAGAAGATT TGTCAAGATTT AACTTTTTG 2280
 GAAGTTGCCA AAGCTCAAGA AGCATGTCAT GGTAAATACTT CAAATAAAGA ACAGTTAACT 2340
 GCTACTAAA CGGAGCAAAA TATAAAAGAT TTTGAGACTT CTGATACATT TTTTCAGACT 2400
 GCAAGTGGGA AAAATATTAG TGTCGCCAAA GAGTCATTAA ATAAAATTGT AAATTCTTT 2460
 GATCAGAAC CAGAAGAATT GCATAACTTT TCCTTAAATT CTGAATTACA TTCTGACATA 2520
 AGAAAAGACA AAATGGACAT TCTAAGTTAT GAGGAAACAG ACATAGTTAA ACACAAAATA 2580
 CTGAAAGAAA GTGTCCCAGT TGGTACTGGA AATCAACTAG TGACCTTCCA GGGACAACCC 2640
 GAACGTGATG AAAAGATCAA AGAACCTACT CTGTTGGGTT TTCATACAGC TAGCGGGAAA 2700
 AAAGTTAAA TTGCAAAGGA ATCTTGGAC AAAGTAAAAA ACCTTTTGAA TGAAAAAAGAG 2760
 CAAGGTACTA GTGAAATCAC CAGTTTGTG CATCAATGGG CAAAGACCCCT AAAGTACAGA 2820
 GAGGCCTGTA AAGACCTTGA ATTGATCATGT GAGACCATTG AGATCACAGC TGCCCCAAAG 2880
 TGTAAAGAAA TGCAGAATTG TCTCAATAAT GATAAAAACC TTGTTCTTAT TGAGACTGTG 2940
 GTGCCACCTA AGCTCTTAAG TGATAATTAA TGTAGACAAA CTGAAAATCT CAAAACATCA 3000
 AAAAGTATCT TTTTGAAAGT TAAAGTACAT GAAAATGTAG AAAAAGAAAC AGCAAAAAGT 3060
 CCTGCAACTT GTTACACAAA TCAGTCCCCT TATTGATCA TTGAAAATTC AGCCTTAGCT 3120
 TTTTACACAA GTTGTAGTAG AAAAACTCT GTGAGTCAGA CTTCAATTACT TGAAGCAAA 3180
 AAATGGCTTA GAGAAGGAAT ATTTGATGGT CAACCCAGAAA GAATAAAATAC TGCAGATTAT 3240
 GTAGGAAATT ATTTGTATGA AAATAATTCA AACAGTACTA TAGCTGAAAA TGACAAAAAT 3300
 CATCTCTCCG AAAAACAAAGA TACTTATTTA AGTAAACAGTA GCATGTCTAA CAGCTATTCC 3360
 TACCAATTCTG ATGAGGTATA TAATGATTC GGATATCTCT CAAAAAATAA ACTTGATTCT 3420
 GGTATTGAGC CAGTATTGAA GAATGTTGAA GATCAAAAAA ACACTAGTT TTCCAAAGTA 3480
 ATATCCAATG TAAAAGATGC AAATGCATAC CCACAAACTG TAAATGAAGA TATTTGCGTT 3540
 GAGGAACCTG TGACTTAGCTC TTACCCCTGC AAAAATAAAA ATGCAGCCAT TAAATTGTCC 3600
 ATATCTAATA GTAATAATT TGAGGTAGGG CCACCTGCAAT TTAGGATAGC CAGTGGTAAA 3660

ATCGTTGTG TTTCACATGA AACAAATTAAA AAAGTGAAG AGACAGTTTC 3720
 AGTAAAGTAA TTAAGGAAAAA CAACGAGAAT AAATCAAAAAA TTTGCCAAAC GAAAATTATG 3780
 GCAGGTTGTT ACGAGGCATT GGATGATTCA GAGGATATTTC TTCATAACTC TCTAGATAAT 3840
 GATGAATGTA GCACGCATTC ACATAAGGTT TTTGCTGACA TTCAGAGTGA AGAAATTAA 3900
 CAACATAACC AAAATATGTC TGGATTGGAG AAAGTTCTA AAATATCACC TTGTGATGTT 3960
 AGTTTGGAAA CTTCAGATAT ATGTAATGTT AGTATAGGGA AGCTTCATAA GTCAGTCTCA 4020
 TCTGCAAATA CTTGTTGGAT TTTTAGCACA GCAAGTGGAA AATCTGTCCA GGTATCAGAT 4080
 GCTTCATTAC AAAACGCAAG ACAAGTGTGTT TCTGAAATAG AAGATAGTAC CAAGCAAGTC 4140
 TTTTCCAAAG TATTGTTAA AAGTAACGAA CATTCAAGACC AGCTCACAAG AGAAGAAAAT 4200
 ACTGCTATAC GTACTCCAGA ACATTTAATA TCCCaaaaAG GCTTTTCATA TAATGTGGTA 4260
 AATTCACTTG CTTCTCTGG ATTTAGTACA GCAAGTGGAA AGCAAGTTTC CATTAGAA 4320
 AGTTCCATTAC ACAAAAGTTAA GGGAGTGTGTT GAGGAATTG ATTTAATCAG AACTGAGCAT 4380
 AGTCTTCACT ATTCCACCTAC GTCTAGACAA AATGTATCAA AAATACTTCC TCGTGTGTT 4440
 AAGAGAAACC CAGAGCACTG TGTAACCTCA GAAATGGAAA AAACCTGCAG TAAAGAATT 4500
 AAATTATCAA ATAACCTAAA TGTTGAAGGT GGTTCTTCAG AAAATAATCA CTCTATTAA 4560
 GTTTCTCCAT ATCTCTCTCA ATTTCAACAA GACAAACAAAG AGTTGGTATT AGGAACCAA 4620
 GTCTCACTTG TTGAGAACAT TCATGTTTG GGAAAAGAAC AGGCTTCACC TAAAACGTA 4680
 AAAATGGAAA TTGGTAAAAC TGAAACTTTT TCTGATGTTTC CTGTGAAAAC AAATATAGAA 4740
 GTTTGTTCTA CTTACTCCAA AGATTCAAGAA AACTACTTTG AAACAGAAAGC AGTAGAAATT 4800
 GCTAAAGCTT TTATGGAAGA TGATGAACCTG ACAGATTCTA AACTGCCAAG TCATGCCACA 4860
 CATTCTCTTT TTACATGTCC CGAAAATGAG GAAATGGTTT TGTCAAATTC AAGAATTGGA 4920
 AAAAGAAGAG GAGAGCCCCCT TATCTTAGTG GGTAAGTGTG CATTAGTACCTT TTTCGTGTG 4980
 CCAATCA 4987

*D /
cont*

(2) INFORMATION FOR SEQ ID NO:113:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 468 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: Genomic DNA
- (ix) FEATURE
 - (A) NAME/KEY: exon
 - (B) LOCATION: 1...468
 - (D) OTHER INFORMATION: Exon 14, nucleotides 12-439 correspond to nucleotides 7236-7663 of complete coding sequence
- (ix) FEATURE
 - (A) NAME/KEY: variable
 - (B) LOCATION: 248
 - (D) OTHER INFORMATION: R at position 248 = A or G.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:113:

CCCCATTGCA GCACAACCTAA GGAACGTCAA GAGATACAGA ATCCAAATT TACCGCACCT 60
 GGTCAAGAAT TTCTGTCTAA ATCTCATTG TATGAACATC TGACTTTGGA AAAATCTTCA 120
 AGCAATTAG CAGTTTCAGG ACATCCATT TATCAAGTTT CTGCTACAAG AAATGAAAAA 180
 ATGAGACACT TGATTACTAC AGGCAGACCA ACCAAAGTCT TTGTTCCACC TTTTAAA 240
 AAATCACRTT TTCACAGAGT TGAACAGTGT GTAGGAATA TTAACCTGGAA GGAAAACAGA 300
 CAAAAGCAA ACATTGATGG ACATGGCTCT GATGATAGTA AAAATAAGAT TAATGACAAT 360
 GAGATTCACT AGTTAACAA AAACAACCTCC AATCAAGCAG CAGCTGTAAC TTTCACAAAG 420
 TGTGAAGAAG AACCTTTAGG TATTGTATGA CAATTGTTGAT GATGAATT 468

(2) INFORMATION FOR SEQ ID NO:114:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 255 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

(A) NAME/KEY: exon

(B) LOCATION: 1...255

(D) OTHER INFORMATION: Exon 22, nucleotides 31-229 correspond to nucleotides 8983-9181 of complete coding sequence

(ix) FEATURE

(A) NAME/KEY: variable

(B) LOCATION: 127

(D) OTHER INFORMATION: R at position 127 = A or G.

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:114:

TTTTTATTCC AATATCTTAA ATGGTCACAG GGTTATTTCA GTGAAGAGCA GTTAAGAGCC 60
TTGAATAATC ACAGGCAAAT GTTGAATGAT AAGAAACAAG CTCAGATCCA GTTGGAAATT 120
AGGAAGRCCA TGGAATCTGC TGAACAAAAG GAACAAGGTT TATCAAGGGA TGTCACAACC 180
GTGTGGAAGT TCGGTATTGT AAGCTATTCA AAAAAAGAAA AAGATTCAAGG TAAGTATGTA 240
AATGCTTGT TTTTA 255

D
D
Cont
(2) INFORMATION FOR SEQ ID NO:115:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 157 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

(A) NAME/KEY: exon

(B) LOCATION: 31...135

(D) OTHER INFORMATION: Exon 2

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:115:

TAAGTGCATT TTGGTCTTCT GTTTTGAGA CTTATTTACC AAGCATTGGA GGAATATCGT 60
AGGTAAAAAT GCCTATTGGA TCCAAAGAGA GGCAACATT TTTGAAATT TTTAAGACAC 120
GCTGCAACAA AGCAGGTATT GACAAATTAA ATATAAC 157

(2) INFORMATION FOR SEQ ID NO:116:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 297 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

(A) NAME/KEY: exon

(B) LOCATION: 21...269

(D) OTHER INFORMATION: Exon 3

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:116:

GGGATTTTTT TTTTAAATAG ATTTAGGACC AATAAGTCTT AATTGGTTG AAGAACCTTC 60
TTCAGAAGCT CCACCCCTATA ATTCTGAACC TGAGAAGAA TCTGAACATA AAAACAACAA 120
TTACGAACCA AACCTATTAA AAACCTCCACA AAGGAAACCA TCTTATAATC AGCTGGCTTC 180

AACTCCAATA ATATTCAAAG AGCAAGGGCT GACTCTGCCG CTGTACCAAT CTCCTGTAAA 240
AGAATTAGAT AAATTCAAAT TAGACTTAGG TAAGTAATGC AATATGGTAG ACTGGGG 297

(2) INFORMATION FOR SEQ ID NO:117:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 159 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: exon
- (B) LOCATION: 26...134
- (D) OTHER INFORMATION: Exon 4

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:117:

D
D
Cmtk
TCACTGAATT ATTGTACTGT TTCAGGAAGG AATGTTCCCA ATAGTAGACA TAAAAGTCTT 60
CGCACAGTGA AAACAAAAAT GGATCAAGCA GATGATGTTT CCTGTCCACT TCTAAATTCT 120
TGTCTTAGTG AAAGGTATGA TGAAGCTATT ATATTAAAA 159

(2) INFORMATION FOR SEQ ID NO:118:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 80 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: exon
- (B) LOCATION: 31...71
- (D) OTHER INFORMATION: Exon 6

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:118:

TTAACAAATT TCCCCTTTT TTACCCCCAG TGGTATGTGG GAGTTGTTT CATAACACCAA 60
AGTTTGTGAA GGTAAATATT 80

(2) INFORMATION FOR SEQ ID NO:119:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 174 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: exon
- (B) LOCATION: 51...165
- (D) OTHER INFORMATION: Exon 7

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:119:

TAATGATCAG GGCATTTCTA TAAAAAATAA ACTATTTCTT TTCCTCCCAG GGTGTCAGA 60
CACCAAAACA TATTCTGAA AGTCTAGGAG CTGAGGTGGA TCCTGATATG TCTTGGTCAA 120
GTTCTTTAGC TACACCACCC ACCCTTAGTT CTACTGTGCT CATAGGTAAT AATA 174

(2) INFORMATION FOR SEQ ID NO:120:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 90 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: exon
- (B) LOCATION: 14...63
- (D) OTHER INFORMATION: Exon 8

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:120:

TTTTATCTTA CAGTCAGAAA TGAAGAAGCA TCTGAAACTG TATTCCTCA TGATACTACT 60
GCTGTAAGTA AATATGACAT TGATTAGACT 90

(2) INFORMATION FOR SEQ ID NO:121:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 142 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: exon
- (B) LOCATION: 20...131
- (D) OTHER INFORMATION: Exon 9

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:121:

TAAACTATAA TTTTGCAGA ATGTGAAAAG CTATTTTCC AATCATGATG AAAGTCTGAA 60
GAAAAATGAT AGATTATCG CTTCTGTGAC AGACAGTGAA AACACAAATC AAAGAGAAGC 120
TGCAAGTCAT GGTAAGTCCT CT 142

(2) INFORMATION FOR SEQ ID NO:122:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 147 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: exon
- (B) LOCATION: 29...124
- (D) OTHER INFORMATION: Exon 12

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:122:

AAAACATATA TGAAATATTT CTTTTAGGA GAACCCTCAA TCAAAAGAAA CTTATTAAAT 60
GAATTTGACA GGATAATAGA AAATCAAGAA AAATCCTTAA AGGCTTCAA AAGCACTCCA 120
GATGGTAAAA TTAGCTTTT ATTATATA 147

(2) INFORMATION FOR SEQ ID NO:123:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 125 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: exon
- (B) LOCATION: 31...100
- (D) OTHER INFORMATION: Exon 13

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:123:

AATATGTAAT ATAAAATAAT TGTTTCCTAG GCACAATAAA AGATCGAAGA TTGTTTATGC 60
ATCATGTTTC TTTAGAGCCG ATTACCTGTG TACCCTTCG GTAAGACATG TTTAAATTTC 120
TCTAA 125

(2) INFORMATION FOR SEQ ID NO:124:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 199 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: exon
- (B) LOCATION: 13...183
- (D) OTHER INFORMATION: Exon 17

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:124:

TTATTTGTTTC AGGGCTCTGT GTGACACTCC AGGTGTGGAT CCAAAGCTTA TTTCTAGAAT 60
TTGGGTTTAT AATCACTATA GATGGATCAT ATGGAAACTG GCAGCTATGG AATGTGCCTT 120
TCCTAAGGAA TTGCTAATA GATGCCTAAG CCCAGAAAGG GTGCTTCTTC AACTAAAATA 180
CAGGCAAGTT TAAAGCATT 199

(2) INFORMATION FOR SEQ ID NO:125:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 380 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: exon
- (B) LOCATION: 19...373
- (D) OTHER INFORMATION: Exon 18

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:125:

TTTTGTTTTC ACTTTTAGAT ATGATACGGA AATTGATAGA AGCAGAACAGAT CGGCTATAAAA 60
AAAGATAATG GAAAGGGATG ACACAGCTGC AAAAACACTT GTTCTCTGTG TTTCTGACAT 120
AATTCATTG AGCGCAAATA TATCTGAAAC TTCTAGCAAT AAAACTAGTA GTGCAGATAAC 180
CCAAAAAGTG GCCATTATTG AACTTACAGA TGGGTGGTAT GCTGTTAAGG CCCAGTTAGA 240

TCCTCCCCCTC TTAGCTGTCT TAAAGAATGG CAGACTGACA GTTGGTCAGA AGATTATTCT 300
TCATGGAGCA GAACTGGTGG GCTCTCCTGA TGCCTGTACA CCTCTTGAGG CCCCAGAAC 360
TCTTATGTTA AAGGTAAATT 380

(2) INFORMATION FOR SEQ ID NO:126:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 208 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: exon
- (B) LOCATION: 30...185
- (D) OTHER INFORMATION: Exon 19

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:126:

TAAATCAATA TATTTATTAA TTTGTCCAGA TTTCTGCTAA CAGTAACCGG CCTGCTCGCT 60
GGTATACCAA ACTTGGATTTC TTTCTGACC CTAGACCTTT TCCTCTGCC 120
TTTCAGTGA TGGAGGAAAT GTTGGTTGTG TTGATGTAAT TATTCAAAGA GCATACCTA 180
TACAGGTATG ATGTATTCTT GAAACTTA 208

D /
cont

(2) INFORMATION FOR SEQ ID NO:127:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 206 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: exon
- (B) LOCATION: 28...172
- (D) OTHER INFORMATION: Exon 20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:127:

TTTGGTGTGT GTAACACATT ATTACAGTGG ATGGAGAAGA CATCATCTGG ATTATACATA 60
TTTCGCAATG AAAGAGAGGA AGAAAAGGAA GCAGCAAAAT ATGTGGAGGC CCAACAAAAG 120
AGACTAGAAG CCTTATTACAC TAAAATTCAG GAGGAATTG AAGAACATGA AGGTAAAATT 180
AGTTATATGG TACACATTGT TATTTC 206

(2) INFORMATION FOR SEQ ID NO:128:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 185 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: exon
- (B) LOCATION: 36...157
- (D) OTHER INFORMATION: Exon 21

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:128:

AGTTTAGTGA ATTAATAATC CTTTGTTT CTTAGAAAAC ACAACAAAAC CATATTTACC 60
ATCACGTGCA CTAACAAGAC AGCAAGTCG TGCTTGCAA GATGGTGCAG AGCTTTATGA 120
AGCAGTGAAG AATGCAGCAG ACCCAGCTTA CCTTGAGGTG AGAGAGTAAG AGGACATATA 180
ATGAG 185

(2) INFORMATION FOR SEQ ID NO:129:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 200 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: exon
- (B) LOCATION: 12...175
- (D) OTHER INFORMATION: Exon 23

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:129:

TCTCCAAACA GTTATACTGA GTATTTGGCG TCCATCATCA GATTATATT CTCTGTTAAC 60
AGAAGGAAAG AGATACAGAA TTTATCATCT TGCAACTTCA AAATCTAAAA GTAAATCTGA 120
AAGAGCTAAC ATACAGTTAG CAGCGACAAA AAAAACTCAG TATCAACAAAC TACCGGTACA 180
AACCTTCAT TGTAATT 200

D
cont
(2) INFORMATION FOR SEQ ID NO:130:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 217 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: exon
- (B) LOCATION: 25...163
- (D) OTHER INFORMATION: Exon 24

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:130:

GAATTTTGT TTTGTTTCT GTAGGTTCA GATGAAATT TATTCAGAT TTACCAGCCA 60
CGGGAGCCCC TTCACTTCAG CAAATTTA GATCCAGACT TTCAGCCATC TTGTTCTGAG 120
GTGGACCTAA TAGGATTGT CGTTCTGTT GTGAAAAAAA CAGGTAATGC ACAATATAGT 180
TAATTTTT TATTGATTCT TTTAAAAAAC ATTGTCT 217

(2) INFORMATION FOR SEQ ID NO:131:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 284 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: exon
- (B) LOCATION: 31...275
- (D) OTHER INFORMATION: Exon 25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:131:

TAACATTCTT TTCTTTTT TCCATTCTAG GACTTGCCCC TTTCGTCTAT TTGTCAGACG 60
AATGTTACAA TTTACTGGCA ATAAAGTTT GGATAGACCT TAATGAGGAC ATTATTAAGC 120
CTCATATGTT AATTGCTGCA AGAACCTCC AGTGGCGACC AGAACCTAAA TCAGGCCTTC 180
TTACTTTATT TGCTGGAGAT TTTCTGCTGT TTTCTGCTAG TCCAAAAGAG GGCCACTTTC 240
AAGAGACATT CAACAAAATG AAAAATACTG TTGAGGTAAG GTTA 284

(2) INFORMATION FOR SEQ ID NO:132:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 210 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: exon
- (B) LOCATION: 31...177
- (D) OTHER INFORMATION: Exon 26

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:132:

ATAAAGCAGC TTTTCCACTT ATTTTCTTAG AATATTGACA TACTTTGCAA TGAAGCAGAA 60
AACAAAGCTTA TGCATATACT GCATGCAAAT GATCCCAAGT GGTCCACCCC AACTAAAGAC 120
TGTACTTCAG GCCCGTACAC TGCTCAAATC ATTCCTGGTA CAGGAAACAA GCTTCTGGTA 180
AGTTAATGTA AACTCAAGGA ATATTATAAG 210

(2) INFORMATION FOR SEQ ID NO:133:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 691 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(ix) FEATURE

- (A) NAME/KEY: exon
- (B) LOCATION: 23...691
- (D) OTHER INFORMATION: Exon 27

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:133:

TACGTTTCA TTTTTTTATC AGATGTCTTC TCCTAATTGT GAGATATATT ATCAAAGTCC 60
TTTATCACCT TGTATGGCCA AAAGGAAGTC TGTTCACCA CCTGTCTCAG CCCAGATGAC 120
TTCAAAGTCT TGTAAAGGGG AGAAAGAGAT TGATGACCAA AAGAACTGCA AAAAGAGAAAG 180
AGCCTTGGAT TTCTTGAGTA GACTGCCCTT ACCTCCACCT GTTAGTCCCA TTTGTACATT 240
TGTTCCTCCG GCTGCACAGA AGGCATTCA GCCACCAAGG AGTTGTGGCA CCAAATACGA 300
AACACCCATA AAGAAAAAAAG AACTGAATTC TCCTCAGATG ACTCCATTAA AAAAATTCAA 360
TGAAATTCT CTTTTGAAAG GTAATTCAAT AGCTGACGAA GAACCTGCA TGATAAATAC 420
CCAAGCTCTT TTGTCTGGTT CAACAGGGAGA AAAACAATT ATATCTGTCA GTGAATCCAC 480
TAGGACTGCT CCCACCAAGTT CAGAAGATTA TCTCAGACTG AAACGACGTT GTACTACATC 540
TCTGATCAA GAACAGGAGA GTTCCCAGGC CAGTACGGAA GAATGTGAGA AAAATAAGCA 600
GGACACAAATT ACAACTAAAA AATATATCTA AGCATTGCA AAGGGCACAA TAAATTATTG 660
ACGCTTAACC TTTCCAGTTT ATAAGACTGG A 691